Tonal bilingualsim: the case of two related Chinese dialects

Wu, J.

Citation

Wu, J. (2015, July 2). *Tonal bilingualsim: the case of two related Chinese dialects*. *LOT dissertation series*. LOT, Utrecht. Retrieved from https://hdl.handle.net/1887/33727

Author: Wu, Junru
Title: Tonal bilingualism : the case of two closely related Chinese dialects
Issue Date: 2015-07-02

Tonal Bilingualism:

The Case of Two Closely Related Chinese Dialects

Published by
LOT phone: +31 30 253 6111
Trans 10
3512 JK Utrecht e-mail: lot@uu.nl
The Netherlands http://www.lotschool.nl

Cover illustration: The SC-JM tonal bilinguals' mind in the author's view, plotted by Junru Wu.

ISBN: 978-94-6093-181-9
NUR 616

Tonal Bilingualism:

The Case of Two Closely Related Chinese Dialects

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 2 Juli 2015
klokke 11.15 uur

door
Junru Wu
geboren te Yizheng, Jiangsu, China
in 1985

Promotores:
  Prof. dr. N.O. (Niels) Schiller
  Prof. dr. V.J.J.P. (Vincent) van Heuven
Co-promotor:
  Dr. Y. (Yiya) Chen
Promotiecommissie:
  Prof. dr. A.F.J. (Ton) Dijkstra
  Prof. dr. C.C. (Claartje) Levelt
  Prof. dr. R.P.E. (Rint) Sybesma
  Dr. M.B. (Martijn) Wieling
  Dr. B. (Björn) Köhnlein

J. Wu's work in this thesis was supported by a PhD Studentship sponsored by Talent and Training China-Netherlands Program, granted by the China Scholarship Council (CSC) and the Netherlands Organisation for Scientific Research (NWO).

Content

2 Tonal Bilingualism: the Case of Two Closely Related Chinese Dialects

1 General Introduction

This thesis investigates the tonal bilingualism involving two closely related Chinese dialects, Standard Chinese (SC) and Jinan Mandarin (JM), a special type of bilingualism. It is special in that it involves (1) two tonal systems, (2) two closely related dialects, and (3) a common logographic writing system.

Focusing on the impact of these factors, this thesis revisits a series of research questions on bilingual lexical process in the context of this special type of bilingualism. Considering interlingual tonal mapping, how do native tonal bilinguals process similar tonal categories from their two tonal systems in lexical access and speech comprehension? Closely related dialects have systematic correspondence between their closely related vocabularies. Is the strength of tonal systematic correspondence in these bilinguals' lexical production affected by the tonal bilinguals' sociolinguistic and cognitive backgrounds? This type of bilingualism involves many etymologically related translation equivalents. How does the tonal similarity between these equivalents affect auditory lexical access? JM shows a significant number of tonal lexical variants. How do the JM speakers handle this tonal pattern variability and store these variants in their mental lexicon? What is the role of interlingual identity in the bilinguals' mental representation and lexical access of these variants? SC and JM use a common logographic writing system; does this affect the bilinguals' automatic visual word recognition? How do they benefit or suffer from this system compared with tonal monolinguals?

In the following sections, after an introduction to SC and JM, the three special aspects of SC-JM tonal bilingualism will be introduced one-by-one, with reviews of relevant research areas, and then brief introductions will be given to each of the main chapters.[i]

1.1 Bilingualism with two tonal systems

Standard Chinese, also frequently referred to as 'Mandarin', 'Mandarin Chinese', 'Putonghua', or just 'Chinese', is a typical tonal language which uses pitch contours to differentiate lexical meanings. It is also one of the most thoroughly examined tonal languages. To maintain consistency, I will only use the term 'Standard Chinese' (SC) in the following sections.

SC is the official language of China (since 1956). Unlike the situation of many European official languages, SC is strictly standardized in its pronunciation. SC speakers can take 'Putonghua Shuiping Ceshi (PSC)' (Putonghua Proficiency Test, since 1994) to see how close their pronunciation is to the standard. Also, SC is mandatorily used in schools within the system of Nine-Year Compulsory Education (except in some administrative units with Minority Compact Communities) and teachers of Chinese in China are required to reach Level 2A (the third highest level) in the PSC test. The national promotion of SC started early and has spread from the urban areas to the rural areas in the past few decades. As a result, most young Chinese throughout China speak SC fluently, many with SC as their first language. Educated SC speakers' accents can be extremely standard. Nevertheless, many Chinese are native bilinguals of SC and their regional Chinese dialects.

Jinan Mandarin (JM) is also a tonal Mandarin dialect spoken by a large population. As a regionally prestigious dialect, JM is spoken in Jinan, the capital of Shandong province in northern China (Qian, 1997). Most JM speakers also speak SC fluently, and the mutual intelligibility between JM and SC is high (Tang & van Heuven, 2009). With the half-century promotion of SC, the demography of bilingualism has changed from JM-dominant to SC-dominant in Jinan. Literacy education used to be carried out mostly in JM but now it is mostly in SC. Although the number of active users of JM is dropping quickly, the society of JM speakers is still relatively large and vital.

Both SC and JM are tonal. The phonological system and pronunciation of SC are standardized according to Beijing Mandarin[ii], which has four citation tones (Brotzman, 1964; Y.R. Chao, 1948; Dreher & Lee, 1968; Fu, 1924), as demonstrated in Figure 1 with examples. JM also has four citation tones (Qian, 1997; Qian & Zhu, 1998), as demonstrated in Figure 2 with examples. Like the other Chinese dialects, both SC and JM use contour tones to distinguish lexical meaning. SC has one famous tone sandhi rule: a low-rising tone (Tone 3) followed by another low-rising tone within the same prosodic phrase is realized as a high-rising tonal contour (similar to Tone2). Additionally, the rising part of the SC low-rising tone is not fully realized in many cases, including when preceding another tone and optionally in fast speech (Lee & Zee, 2003). JM tone sandhi is more complex. The rules have been under some discussion and the reported pattern is ambiguous (Qian, 1997). The details of some aspects and the corresponding causes will be introduced and investigated in the following chapters. The dominant rules assumed in this thesis are as follows:

Rising+High-falling -> Low+High-falling
Rising+High-level -> Low+High-level
Low-falling+Low-falling -> Rising+Low-falling[iii]
Rising+neutral -> Low-falling+Low
High-falling+neutral -> Rising+High
High-level+neutral -> Low+High
Low-falling+neutral -> High-level+neutral

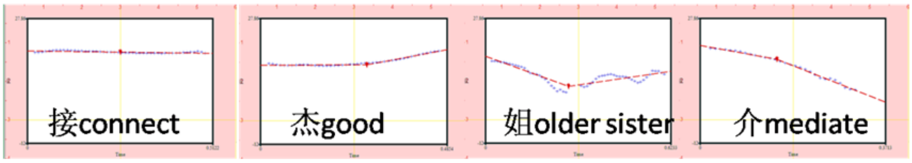


Figure 1. *Pitch contours on the rhymes of a SC tonal minimal set with the same segmental structure /tɕiɛ/ [from left to right, 1: high-level, 2: high-rising, 3: low-rising (dip); 4: falling]. It is pronounced by a young female SC native speaker with PSC 1b and plotted with a custom-made piecewise regression function in Praat (Boersma & Weenink, 2014). Examples of SC pitch contours plotted with early recordings can be found by Brotzman (1964), Dreher & Lee (1968), and Fu (1924).*

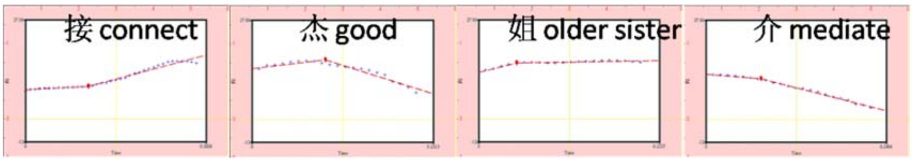


Figure 2. *Pitch contours on the rhymes of a JM tonal minimal set with the same segmental structure /tɕiɛ/*[iv] *[from left to right, 1: rising, 2: high-falling, 3: high-level; 4: low-falling], pronounced by an old male JM speaker recorded by Qian (1998), plotted with a custom-made piecewise regression function in Praat (Boersma & Weenink, 2014).*

For both JM and SC, morphemes carrying neutral tones lose the contrast of their citation tonal forms, and the F0 realization of the neutral tone varies, depending on the preceding citation tone. Nevertheless, JM neutral tone sandhi is different from the case of SC, in that JM morphemes preceding neutral tones are realized as different sandhi forms from their citation forms, while in the same position SC morphemes largely maintain the citation form.

The bilingualism of SC and JM involves two systems of lexical tones. In the following four subsections, we will review the uniqueness of tonal languages and the results of bilingual studies involving tonal languages.[v]

1.1.1 The uniqueness of tonal languages

The uniqueness of tonal language processing has been investigated extensively, especially in speech perception and lexical access.

As with speech perception, categorical perception of native tones is supported by recent neurological studies, in which the mismatch negativity (MMN) is larger across categories than within categories (Chandrasekaran, Krishnan, & Gandour, 2009; Xi, Zhang, Shu, Zhang, & Li, 2010). However, behavioral studies showed that, compared with segmental phonemes, native perception of Chinese tones is ‘quasi-categorical’, neither as categorical as that of consonants nor as continuous as that of vowels (Hallé, Chang, & Best, 2004). The retrieval of tonal information, compared with segmental information seems to be slower (Ye & Connine, 1999; Q. Zhang & Damian, 2009; Q. Zhang & Zhu, 2011) and involves different neuronal networks (Liang & van Heuven, 2004).

Experience with tonal languages changes tone-related perception and the corresponding neurological activities. Faced with between-category contrasts in odd-ball paradigms, tonal language speakers showed larger mismatch negativity (MMN) responses than non-tonal language speakers (Chandrasekaran, Krishnan, & Gandour, 2007; Chandrasekaran et al., 2009) to the mismatch. The tonal language speakers behave more similarly to musicians than non-musicians who do not speak a tonal language, showing a domain-general advantage (Chandrasekaran et al., 2009); Tonal language speakers have general advantage in discriminating musical tones and are less affected by pitch perturbation (Ning, Shih, & Loucks, 2014). When making a phonetic decision, tonal language speakers and non-tonal language speakers showed differences in selective attention: speakers of tonal language (SC) process consonants and tones in a combined manner, but non-tonal language (English) speakers can separate them in the process of speech perception (M. Lin & Francis, 2014). Experience with tonal language not only modifies the cortical distribution of stimulus-dependent activation (Hsieh, Gandour, Wong, & Hutchins, 2001; Krishnan, Gandour, Ananthakrishnan, & Vijayaraghavan, 2014), but also strengthens the correlation between cortical pitch response (CRP) components and pitch acceleration (Krishnan et al., 2014).

As with lexical access, lexical tones are also processed differently from segmental phonemes. The overlap of SC tones alone induced no facilitatory priming effect in implicit priming (J. Y. Chen, Chen, & Dell, 2002), differing from the classical priming effect introduced by segmental overlap. Nor does sharing both surface tones and segments cause facilitatory priming (Y. Chen, Shen, & Schiller, 2011). However, tonal sharing (tonemes or overt tonal realizations) accompanied with segmental sharing introduced phonological facilitation in two picture-word interference experiments (Nixon, Chen, & Schiller, 2014), similar to the case of segmental sharing. Lexical adaptation also seems to work similarly in tones (Mitterer, Chen, & Zhou, 2011) and consonants (McQueen, Cutler, & Norris, 2006). Moreover, eye-tracking evidence support that, in constraining activation, tonal and segmental accesses are concurrent and play comparable roles (Malins & Joanisse, 2010).

1.1.2 Perceptual learning of non-native lexical tones

Recent studies investigated the perceptual learning of non-native lexical tones. On the one hand, native experience with tonal languages interacts with the brain networks. Learning non-native tones changes the brain network (Yang, Gates, Molenaar, & Li, 2014) and strengthens the brain response to within-category tonal contrast in right STG (Zinszer, Chen, Wu, Shu, & Li, 2014). Also, individual differences in white matter pathways can predict the learning success of a tonal language (SC) (Qi, Han, Garel, San Chen, & Gabrieli, 2014).

On the other hand, experience with tonal languages influences perceptual learning of a new tonal language. It is found that, in perceptual learning, non-native tones (Cantonese lexical tones) triggered different changes in the identification and perceptual space of speakers with native tonal experience (SC) versus speakers without native tonal experience (English) (Francis, Ciocca, Ma, & Fenn, 2008). However, findings on whether L1 lexical tone experience necessarily helps the learning of L2 tonal perception are inconsistent. For instance, while Mandarin speakers showed greater advantages over English speakers in distinguishing Thai tones after training (Wayland & Guion, 2004), Hmong learners performed surprisingly worse than English learners in perceiving Chinese tones (Wang, 2006).

1.1.3 Interlingual tonal perception

Previous studies involving two tonal languages have focused on interlingual tonal perception. After early focus on the matching of lexical tone and lexical stress in non-native production (Y. R. Chao, 1980; Cheng, 1967; Cheung, 2008), research attention was drawn to tonal perception across different tonal languages (Reid et al., 2014; So & Best, 2010; Xujin Zhang, Samuel, & Liu, 2012). So and colleagues' studies focus on naïve listeners and were discussed under the Perceptual Assimilation Model (PAM) (C. T. Best, 1995) which claims that naïve listeners perceive non-native phones according to the closest L1 phonemes (if they exist). However, recent findings suggest that tonal assimilation also needs to take phonetic and even acoustic similarity into consideration (Reid et al., 2014). The other two popular models for interlingual perception are the SLM model (J. E. Flege, 1995; James Emil Flege, MacKay, & Piske, 2002), which focuses more on L2 learning, and the native-language magnet (NLM) model (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), which focuses more on perception within the L1 category. I have not found studies of interlingual tonal perception discussed under these two models. Zhang and colleagues' research showed that, although they have early and rich experience using SC, native speakers of Cantonese are clearly influenced by Cantonese when handling SC contrasts in acoustic, perceptual, and lexical processing. They were also the first to find that the tonal effects are of a similar pattern but smaller than segmental effects (Xujin Zhang et al., 2012). A recent study provided some new insights into the role of pitch in bilingual tonal perception: tonal and non-tonal language users may attend to the same pitch movements in different ways (Braun & Johnson, 2011). Despite these studies, little is known about the interlingual tonal perception of real tonal bilinguals. B. Chen reported the development of Tai-accented Chinese lexical production in child second language learners (who later become tonal bilingual adults, like a significant portion of that linguistic society) and associated his findings (including the change of tonal mapping) with macroscopic language contact and language evolution (B. Chen, 1996). Nevertheless, the mental process of interlingual tonal perception by real tonal bilinguals remains largely unknown.

The above-mentioned two sets of studies focus on naïve listeners or beginning learners of tonal languages. Real tonal bilinguals, people who already speak two tonal languages/dialects, however, receive less attention. A large number of residents in China speak SC and at least one tonal dialect or language natively, which can be another Mandarin dialect (e.g. Jinan Mandarin), a more distant Chinese dialect (e.g. Cantonese), or another tonal language spoken in China (e.g. Tai dialects spoken in southwest China[vi]). Two previous studies involving Cantonese-SC bilinguals showed some unique attributes of these tonal bilinguals in speech processing. Children who are bilingual in Cantonese and SC showed more advanced tonal awareness compared with tonally monolingual SC speaking children (X. Chen et al., 2004). Also, in a recent study, the correlation between naming performance and left inferior parietal lobule (LIPL) volume was more prominent for Mandarin-Cantonese tonal bilinguals than for Cantonese-English bilinguals (Abutalebi, Canini, Della Rosa, Green, & Weekes, 2014). Nevertheless, the relations of the two vocabularies and two tonal systems need to be taken into consideration and the role of tone in bilingual speech perception and lexical access needs more investigation. SC-JM bilingualism makes a good test case.

1.1.4 SC in bilingual studies

It is not uncommon to include tonal languages in bilingual studies. Standard Chinese, as one of the world's most widely spoken languages, was included in bilingual experiments to investigate different research questions.

For instance, proficient late SC-English bilinguals showed onset priming in English but syllabic priming in Chinese in word-naming, indicating that they use different phonological units (syllables and phonemes) to fill the metrical frame according to the language mode (Verdonschot, Nakayama, Zhang, Tamaoka, & Schiller, 2013). However, acoustic analysis of late SC-English bilinguals' English production indicates that the rhythm of the interlanguage (Chinese accented English) needs to be interpreted with both SC and English rhythmic units under consideration (H. Lin & Wang, 2008).

Consistent with the results found between other languages, the SC findings support parallel bilingual lexical activation. Researchers have found interlingual semantic priming (Keatley, Spinks, & De Gelder, 1994), translation priming (H.-C. Chen & Ng, 1989), and cross-language identity effects (Taomei Guo & Peng, 2006) between SC and English. Also, late SC-English bilinguals are sensitive to concealed relations of Chinese characters when performing lexical comprehension tasks in English, indicating unconscious translation to the native language (Thierry & Wu, 2004; Y.J. Wu & G. Thierry, 2010; Yan Jing Wu & Guillaume Thierry, 2010; Y. J. Wu & Thierry, 2011, 2012).

As with bilingual language and executive control, SC learners of English were tested for the neural correlates of global and local language-switching costs (T. Guo, Liu, Misra, & Kroll, 2011; Prior & Gollan, 2011). Comparing SC-English bilinguals with English monolinguals, the bilinguals showed executive advantages (Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011), as has been found with bilinguals of other languages (Bialystok, Craik, & Luk, 2008). Moreover, in SC-English bilinguals it was found that both age of acquisition, (AOA) and L2 proficiency affect the type of executive advantages and the corresponding attentional networks (Tao et al., 2011).

As with L2 acquisition, SC learners of English with different individual language histories and from different linguistic communities showed different L2 lexical categorization, indicating the importance of these two factors (Zinszer, Malt, Ameel, & Li, 2014). SC L2 learners of English were tested for cross-language transfer effects under syntactical violation, and the absence of such effects support the unified model of language acquisition (Tuninetti, Warren, & Tokowicz, 2014). Balanced and unbalanced SC-English bilinguals showed different neural correlates with the increasing load on phonological working memory (PWM), indicating the importance of PWM in language attainment (Chee, Soon, Lee, & Pallier, 2004).

SC was also used to study L2 reading of different types of scripts. Alphabetic language speakers learning SC as L2 showed different neural correlates of regularity effects in Chinese reading, compared with SC native speakers (Zhao et al., 2012).

1.2 Bilingualism of closely related dialects

Most previous bilingual studies focus on the bilingualism of different standard languages (e.g. English-Dutch, English-Spanish). Nevertheless, the bilingualism of closely related dialects (or languages), although very common, only received minimal attention. The limited studies are primarily present in machine translation (Xiaoheng Zhang, 1998), automatic (machine) speech recognition (Sproat et al., 2004), and the impact of local dialects on foreign language education (Feng & Adamson, 2014).

Although recently a few studies have investigated the speech perception (Sumner & Samuel, 2009; Xujin Zhang et al., 2012), lexical access (Xujin Zhang et al., 2012), and interlingual intelligibility (Gooskens, Kürschner, & van Bezooijen, 2011; Tang & van Heuven, 2009) of closely related languages and dialects, two aspects of interlingual alignments of the related phonological systems and their impact on the bilingual mental processing of speech recognition and production clearly invite more investigation. Different from two remote languages, two closely related languages or dialects not only show phonological similarity in their basic sound inventories, but also share a large number of etymologically related translation equivalents (including cognates and loan words). As a result, both phonological similarity and systematic correspondence are prevalent within the etymologically related translation equivalents.

In the present case, JM and SC both are northern Mandarin dialects. They have almost identical segmental inventories and similar tonal inventories (both with high-level, falling, and rising tones). Because of their historical relation, their etymologically related translation equivalents are usually segmentally identical and only vary in tonal similarity. To make it more interesting, the tones of the etymologically related words ‘correspond systematically’, according to the definition of systematic correspondence (Dyen, 1963; Meillet & Ford, 1967), so that the tonal categories of monosyllabic JM words are, to a large extent, predictable from the tonal categories of their SC translation equivalents. The introduction of Chapter 3 will provide more detailed descriptions of the phonological similarity and systematic correspondence in SC and JM with examples.

1.3 Bilingualism involving a common logographic writing system

Bilingual reading has been investigated in depth. Interlingual activation happens in automatic visual word recognition, either when the involved language uses the same type of orthography (Dyer, 1971; Preston & Lambert, 1969) or not (H.-C. Chen & Ho, 1986; Fang, Tzeng, & Alva, 1981; Kiyak, 1982). The within-language effect is usually greater than the between-language effect (H.-C. Chen & Ho, 1986; Fang et al., 1981; Preston & Lambert, 1969). However, a few studies have shown beginning L2 learners may experience between-language interference equally to (Kiyak, 1982), or even more than (H.-C. Chen & Ho, 1986) within-language interference when L2 is the response language in Stroop tests. The type of script affects visual word recognition. Users of alphabetic writing systems as learners of Chinese recruit different brain networks compared with native Chinese readers in reading phonological regular and irregular Chinese characters (Zhao et al., 2012). Between-language Stroop effects are also affected by the similarity of the scripts (Fang et al., 1981; Van Heuven, Conklin, Coderre, Guo, & Dijkstra, 2011).

However, bilingualism involving a common logographic writing system is little studied. Although this situation sounds very unlikely, it is quite common across Chinese dialects. For instance, in the present SC-JM case, the bilingual lexicon is teaming with etymologically related translation equivalents; and the bilinguals use the same Chinese characters for these pairs of SC-JM translation equivalents, such as the same Chinese character 蓝 for both the SC ‘blue’ /lan(High-rising) / and the JM ‘blue’ /lan(High-falling)/. Which information is activated in bilingual visual word recognition with these Chinese characters? Do tonal bilinguals read Chinese characters in the same way as tonal monolinguals?

1.4 Interlingual tonal mapping

As mentioned in the first section, Interlingual tonal perception has been investigated in naïve non-native listeners (So, 2010; Reid 2014). In addition to this, two-to-one interlingual mapping of phonemes is not a rare phenomenon in bilingualism and has been investigated intensively. One of the most famous cases is the mapping of English /r/ and /l/ with the Japanese apico-alveolar tap /ɾ/, with the Japanese /ɾ/ perceptually more similar to /l/ (Catherine T. Best & Strange, 1992; Miyawaki et al., 1975). Such a phenomenon can be approached from different directions, such as perceptual mapping (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Iverson et al., 2003) and interlingual lexical competition (Cutler, Weber, & Otake, 2006).

However, how native tonal bilinguals process similar tonal categories from their two tonal systems remains unclear. Chapter 2 looks into the special situation of interlingual tonal mapping by SC-JM tonal bilinguals, namely the two-to-one interlingual tonal mapping of SC high-rising and low-rising tones with the JM rising tone. This two-to-one tonal mapping is investigated in the final position of disyllabic JM words and the corresponding SC pseudo-words. The native acoustic distributions of these three rising tones, as well as the interlingual identification of the JM rising tone by SC monolinguals, are compared and modeled with Generalized Additive Modeling (GAM). Then a semantic priming experiment investigates whether and how SC rising tones activate JM words in native bilinguals’ mental lexicon, and to what extent the interlingual category-goodness keeps its influence on speech comprehension.

1.5 Systematic correspondence between etymologically related translation equivalents

As mentioned in the second section, bilingualism involving related dialects has many etymologically related translation equivalents, which demonstrate systematical correspondence and phonological similarity. These two phonological alignments between two vocabularies have received intensive attention from historical linguistics since the early studies of historical comparison (Dyen, 1963; Meillet & Ford, 1967). However, they have received relatively less attention from psycholinguistics and speech engineering. These two aspects are mainly studied in Chapter 3 and 4.

Chapter 3 carries out an explorative study, trying to statistically model the pitch between JM words from the tonal relation of their SC translation equivalents. This study is primarily based on the rule of systematic correspondence, which predicts that, if two words are from the same phonological category in one language, their etymologically related translation equivalents in the corresponding language are also more likely to come from the same phonological category, although the involved phonological categories from the two languages may sound very different.

Besides statistically verifying the effects of systematic correspondence, the present study also investigates how different sociolinguistic and cognitive backgrounds affect the strength of systematic correspondence in bilingual individuals. Individual backgrounds are incorporated into the modeling of systematic correspondence. Also, age-dependent and age-independent effects are statistically separated to clarify the general sources of individual variability in the bilingual society. The results may be interesting for the linguistic studies of language contact and language evolution, as well as for the practice of speech engineering.

1.6 Phonological similarity and lexical variants of etymologically related translation equivalents

Interlingual similarity of phonological categories is common in bilingualism and we have given a brief introduction to the SC-JM interlingual two-to-one tonal mapping in a previous section (Interlingual Tonal Mapping) and a more detailed investigation will be given in Chapter 2. Nevertheless, only closely related languages (or dialects) or languages in close contact have a lot of phonologically similar translation equivalents. Phonological identity is an extreme form of phonological similarity. In the current section, a general introduction is given for Chapters 4 and 6, which investigates the effects of phonological similarity in the bilingual mental lexicon and lexical access under the assumption of an integrated bilingual lexicon (Kroll, Bobb, & Wodniecka, 2006).

Etymologically related translation equivalents include both cognates and loan words, which are difficult to distinguish in closely related dialects. Using 'cognates' to refer to all etymologically related translation equivalents, psycholinguists have found 'cognate facilitation effect' in many different tasks and conditions. For instance, 'cognates' are processed faster in both production (Costa, Caramazza, & Sebastian-Galles, 2000; Hoshino & Kroll, 2008) and visual word recognition (Brenders, van Hell, & Dijkstra, 2011; Bultena, Dijkstra, & van Hell, 2012; Dijkstra, Grainger, & Van Heuven, 1999). 'Cognate facilitation' comes both from the sharing of orthography and phonology. There have been many studies trying to tear them apart, where the orthographic effect is more robust than the phonological effect

(Dijkstra et al., 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Duyck, Assche, Drieghe, & Hartsuiker, 2007; Lemhöfer & Dijkstra, 2004).

SC-JM provides an ideal case to test the tonal 'cognate effect' aurally, in that the translation equivalents are mostly etymologically related. Also these etymologically related translation equivalents are mostly identical in segmental structure and orthography, and vary only in tonal similarity. Thus, the phonological similarity between a pair of SC-JM etymologically related translation equivalents only depends on the tonal similarity.

Chapter 4 investigates the role of tone in the auditory lexical access of etymologically related translation equivalents. SC-JM tonal bilinguals and SC monolinguals make auditory lexical decisions about tone-identical and tone-non-identical translation equivalents, with the tonal similarity of each pair assessed afterwards. The naming latencies are measured and compared across different types of translation equivalents, as well as between tonal bilinguals and tonal monolinguals, in order to investigate the lexical representation of these SC-JM translation equivalents.

Possibly due to lack of standardization and heavy language contact, many JM words in speech production show tonal pattern variability, a variability between tonal lexical variants. Chapter 5 investigates the roles of three different types of tonal variability in JM auditory lexical access, with a focus on the tonal pattern variability. This is a type of variability with lexical variants of the same word carrying distinctive tonal patterns, which are lexically non-contrastive but potentially contrastive in other words. This is a case similar to the vowel variability of /a:/ and /ei/ in the British and American pronunciation of 'tomato'. What differs is that the tonal pattern variability exists within the same language and used by the same speaker. The form priming results show that tonal pattern variants induced a similar but smaller facilitation effect compared with the acoustic identity and the within-category variation, which is different from the inhibition effect in the lexically contrastive condition. Thus, tonal patterns may have representative status but can converge in a lexically specific way in lexical access.

In Chapter 6, the storage of tonal pattern variants is investigated with the tonal bilingualism under consideration. Taking the phenomena discussed in Chapter 4 and Chapter 5 together, some JM words in JM-only mode are produced with different lexical variants with distinctive tonal patterns and one of the variants is identical to the SC translation equivalent (variant_id). Moreover, in a corpus including different speakers, the word-wise probability of variant_id varies across words, ranging from 0 to 1. Do the bilinguals who only produce the variant_id also store the non-identical variant in their mental lexicon? In previous studies, variant frequency effect is used to support the separate lexical representations of the flapping and non-flapping variants of American English consonant /t/ (Connine, Ranbom, & Patterson, 2008). In Chapter 6, the variant frequency effects on the naming latencies of JM words are tested to investigate the lexical representation of the interlingual identical and non-identical variants.

1.7 Automatic bilingual phonological activation from the common written forms

As mentioned in the third section, in earlier studies, bilingualism involving the same type of orthography has been compared with bilingualism involving different types of orthographies (Fang et al., 1981; Van Heuven et al., 2011). However, the usage of exactly the same written forms for translation equivalents in bilingualism is little studied.

Regarding the visual word recognition of Chinese, a logographic writing system, phonological effects have been investigated in depth in previous studies. It was found very early that phonological information is retrieved in Chinese character recognition (Tzeng, Hung, & Wang, 1977). A meta-study shows that the brain regions involved in phonological processing differ between Chinese characters and written alphabetic words (Tan, Laird, Li, & Fox, 2005). Behavioral studies found that, different from alphabetic writing systems, Chinese characters can be identified without mediation from phonemic processes (Perfetti & Zhang, 1991). Another study found that evidence for phonological mediation is only restricted to words with relatively thin homophone density (Tan & Perfetti, 1997). Although the phonological activation can happen automatically and early in visual word identification (Perfetti & Zhang, 1991, 1995), it mostly occurs postlexically (Perfetti & Zhang, 1991; Zhou & Marslen-Wilson, 2000). Various priming experiments found that the existence of phonetic radicals affects the speed and the role of phonological activation in Chinese visual word identification (N. Wu, Zhou, & Shu, 1999; Zhou & Marslen-Wilson, 1999b) and the phonological activation is largely mediated by the phonetic radicals (Zhou & Marslen-Wilson, 1999a). Nevertheless, these studies did not take into consideration that a large proportion of Chinese readers use exactly the same characters for translation equivalents from different Chinese dialects. With these bilinguals, which phonological representations are activated by the common Chinese character, and whether the different tonal representations from both dialects are automatically activated via the same Chinese character needs more investigation.

In Chapter 7, the automatic phonological activation of Chinese characters is tested in Stroop experiments. Similar to earlier Chinese Stroop experiments (C. Li, Lin, Wang, & Jiang, 2013; Spinks, Liu, Perfetti, & Tan, 2000), participants name the ink color of different words written in different colors, and the naming latencies and accuracies are measured. However, different from previous studies, the SC-JM tonal bilinguals and SC tonal monolinguals are tested separately and in different dialect modes. The within-dialect and between-dialect phonological sharing between the character and color names is manipulated, in order to assess the tonal and segmental effects. The Stroop facilitation and interference are measured both on the tonal bilinguals and tonal monolinguals, in order to investigate the bilingual effect. The results are compared with previous findings and discussed in light of tone-specific bilingual attention control.

References

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., & Weekes, B. S. (2014). The neuroprotective effects of bilingualism upon the inferior parietal lobule: A Structural Neuroimaging Study in Aging Chinese Bilinguals. *Journal of Neurolinguistics, 33*(0), 3-13. doi: 10.1016/j.jneuroling.2014.09.008

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r. *Journal of Phonetics, 32*(2), 233-250.

Benedict, P. K. (1942). Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *American Anthropologist, 44*(4), 576-601.

Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. *Speech perception and linguistic experience: Issues in cross-language research*, 171-204.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859. doi: 10.1037/0278-7393.34.4.859

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] Version 5.3.82 retrieved 26 July 2014 from http://www.praat.org/.

Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics, 39*(4), 585-594.

Brenders, P., van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of experimental child psychology, 109*(4), 383-396. doi: 10.1016/j.jecp.2011.03.012

Brotzman, R. (1964). Progress report on Mandarin tone study: Ohio State Univ., Columbus. Research Foundation.

Bultena, S., Dijkstra, T., & van Hell, J. G. (2012). Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes*(ahead-of-print), 1-28. doi: 10.1080/01690965.2012.718353

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain research, 1128*, 148-156. doi: 10.1016/j.brainres.2006.10.064

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*(1), 1-9. doi: 10.1016/j.bandl.2008.02.001

Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese* (Vol. 1): Harvard University Press.

Chao, Y. R. (1980). Chinese Tones and English Stress. In L. Waugh & C. v. Schooneveld (Eds.), *The melody of language* (pp. 41-44). Baltimore: University Park Press.

Chee, M. W. L., Soon, C. S., Lee, H. L., & Pallier, C. (2004). Left insula activation: A marker for language attainment in bilinguals. *Proceedings of the National Academy of Sciences of the United States of America, 101*(42), 15265.

Chen, B. (1996). *On Language Contact and Language Union (Lun Yuyan Jiechu Yu Yuyan Lianmeng 论语言接触与语言联盟)*: Language and Literature Press (Yuwen Chubanshe 语文出版社).

Chen, H.-C., & Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(3), 397. doi: 10.1037/0278-7393.12.3.397

Chen, H.-C., & Ng, M.-L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition, 17*(4), 454-462.

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751-781. doi: 10.1006/jmla.2001.2825

Chen, X., Anderson, R. C., Li, W., Hao, M., Wu, X., & Shu, H. (2004). Phonological Awareness of Bilingual and Monolingual Chinese Children. *Journal of Educational Psychology, 96*(1), 142-151. doi: 10.1037/0022-0663.96.1.142s

Chen, Y., Shen, R., & Schiller, N. O. (2011). *Representation of allophonic tone sandhi variants.* Paper presented at the Proceedings of Psycholinguistics Representation of Tone. Satellite Workshop to ICPhS, Hongkong.

Cheng, C. C. (1967). English Stresses and Chinese Tones in Chinese Sentences (Vol. POLA-2-1): California University, Berkeley, Phonology Laboratory.

Cheung, W. H. Y. (2008). Span of High Tones in Hong Kong English. *HKBU Papers in Applied Language Studies, 12*, 19-46.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics, 70*(3), 403-411.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283. doi: 10.1037/0278-7393.26.5.1283

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284. doi: 10.1016/j.wocn.2005.06.002

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301. doi: 10.1016/j.jml.2009.12.003

Dreher, J. J., & Lee, P.-c. e. (1968). Instrumental investigation of single and paired Mandarin tonemes. *Monumenta serica*, 343-373.

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 663. doi: 10.1037/0278-7393.33.4.663

Dyen, I. (1963). Why phonetic change is regular. *Language, 39*(4), 631-637. doi: 10.2307/411958

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior, 10*(3), 297-302. doi: 10.1016/S0022-5371(71)80057-9

Fang, S.-P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition, 9*(6), 609-617. doi: 10.3758/BF03202355

Feng, A., & Adamson, B. (Eds.). (2014). *Trilingualism in Education in China: Models and Challenges* (Vol. 12): Springer.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.

Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics, 23*(04), 567-598.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36*(2), 268-294.

Fu, L. (1924). *Experimental report of 4 tones in Mandarin*: Shanghai Qunyi Publishing House.

Gooskens, C., Kürschner, S., & van Bezooijen, R. (2011). Intelligibility of High and Low German to speakers of Dutch. *Dialectologia, Special issue, 2*, 35-63.

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *Neuroimage*. doi: 10.1016/j.neuroimage.2011.03.049

Guo, T., & Peng, D. (2006). Event-related potential evidence for parallel activation of two languages in bilingual speech production. *NeuroReport, 17*(17), 1757-1760.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*(3), 395-421. doi: 10.1016/S0095-4470(03)00016-0

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106*(1), 501-511. doi: 10.1016/j.cognition.2007.02.001

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language, 76*(3), 227-252. doi: 10.1006/brln.2000.2382

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57. doi: 10.1016/S0010-0277(02)00198-1

Keatley, C., Spinks, J., & De Gelder, B. (1994). Asymmetrical semantic facilitation between languages. *Memory & Cognition, 22*, 70-84.

Kiyak, H. A. (1982). Interlingual interference in naming color words. *Journal of Cross-Cultural Psychology, 13*(1), 125-135. doi: 10.1177/0022022182131011

Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2014). Language experience enhances early cortical pitch-dependent responses. *Journal of Neurolinguistics, 33*(0), 128-148. doi: 10.1016/j.jneuroling.2014.08.002

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9*(02), 119-135. doi: 10.1017/S1366728906002483

Kuhl, P. K. (1991). Human adults and human infants show a' perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50*(2), 93-107.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606.

Lee, W.-S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association, 33*(1). doi: 10.1017/S0025100303001208

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32*(4), 533-550. doi: 10.3758/BF03195845

Li, C., Lin, C., Wang, M., & Jiang, N. (2013). The activation of segmental and tonal information in visual word recognition. *Psychonomic bulletin & review*. doi: 10.3758/s13423-013-0395-2

Li, F.-k. (1973). Languages and Dialects of China. *Journal of Chinese Linguistics, 1*(1), 1-13.

Liang, J., & van Heuven, V. J. (2004). Evidence for separate tonal and segmental tiers in the lexical specification of words: A case study of a brain-damaged Chinese speaker. *Brain and Language, 91*(3), 282-293. doi: 10.1016/j.bandl.2004.03.006

Lin, H., & Wang, Q. (2008). *Interlanguage rhythm in the English production of Mandarin speakers*. Paper presented at the Proceedings of the 8th Phonetic Conference of China.

Lin, M., & Francis, A. L. (2014). Effects of language experience and expectations on attention to consonants and tones in English and Mandarin Chinese. *The Journal of the Acoustical Society of America, 136*(5), 2827-2838.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62*(4), 407-420. doi: 10.1016/j.jml.2010.02.004

Matisoff, J. A. (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126. doi: 10.1207/s15516709cog0000_79

Meillet, A., & Ford, G. B. (1967). *The comparative method in historical linguistics*. Paris: H. Champion.

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: evidence from a learning paradigm. *Cognitive Science, 35*(1), 184-197. doi: 10.1111/j.1551-6709.2010.01140.x

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*(5), 331-340. doi: 10.1111/j.1551-6709.2010.01140.x

Ning, L.-H., Shih, C., & Loucks, T. M. (2014). Mandarin tone learning in L2 adults: A test of perceptual and sensorimotor contributions. *Speech Communication, 63*, 55-69. doi: 10.1016/j.specom.2014.05.001

Nixon, J. S., Chen, Y., & Schiller, N. O. (2014). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. *Language, Cognition and Neuroscience, 30*(5), 491-505. doi: 10.1080/23273798.2014.942326

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633. doi: 10.1037/0278-7393.17.4.633

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24. doi: 10.1037/0278-7393.21.1.24

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 295-301. doi: 10.1016/S0022-5371(69)80079-4

Prior, A., & Gollan, T. H. (2011). Good Language-Switchers are Good Task-Switchers: Evidence from Spanish-English and Mandarin- English Bilinguals. *Journal of the International Neuropsychological Society, 17*(4), 682. doi: 10.1017/S1355617711000580

Qi, Z., Han, M., Garel, K., San Chen, E., & Gabrieli, J. D. E. (2014). White-matter structure in the right hemisphere predicts Mandarin Chinese learning success. *Journal of Neurolinguistics, 33*(0), 14-28. doi: 10.1016/j.jneuroling.2014.08.004

Qian, Z.-Y. (Ed.). (1997). *'Jinan Fangyan Cidian' (Jinan Dialect Dictionary)*. Nanjing: Jiangsu Education Press.

Qian, Z.-Y., & Zhu, G.-Q. (1998). 'Jinanhua Yindang' (The Sound System of Jinan Dialect): Shanghai Educational Publishing House, Shanghai.

Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Rattanasone, N. X. (2014). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Attention, Perception, & Psychophysics*, 1-21.

So, C. K., & Best, C. T. (2010). Cross-language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences. *Language and speech, 53*(2), 273.

Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: The role of phonological information. *Cognition, 76*(1), B1-B11. doi: 10.1016/S0010-0277(00)00072-X

Sproat, R., Zheng, F., Gu, L., Jurafsky, D., Shafran, I., Li, J. (2004). Dialectal chinese speech recognition: Final report. doi:10.1.1.133.7189

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501. doi: 10.1016/j.jml.2009.01.001

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25*(1), 83-91. doi: 10.1002/hbm.20134

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*(1), 41-57. doi: 10.1006/jmla.1997.2508

Tang, C., & van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua, 119*(5), 709-732. doi: 10.1016/j.lingua.2008.10.001

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00123

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport, 15*(10), 1555-1558.

Tuninetti, A., Warren, T., & Tokowicz, N. (2014). Cue strength in second-language processing: An eye-tracking study. *The Quarterly Journal of Experimental Psychology*(ahead-of-print), 1-17. doi: 10.1080/17470218.2014.961934

Tzeng, O. J., Hung, D. L., & Wang, W. S. Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory, 3*(6), 621-630. doi: 10.1037/0278-7393.3.6.621

Van Heuven, W. J. B., Conklin, K., Coderre, E. L., Guo, T., & Dijkstra, T. (2011). The influence of cross-language similarity on within-and between-language Stroop effects in trilinguals. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00374

Verdonschot, R. G., Nakayama, M., Zhang, Q., Tamaoka, K., & Schiller, N. O. (2013). The Proximate Phonological Unit of Chinese-English Bilinguals: Proficiency Matters. *PLoS ONE, 8*(4), e61454.

Wang, X. (2006). Perception of L2 tones: L1 lexical tone experience may not help. *Speech Prosody 2006*.

Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*(4), 681-712.

Wu, N., Zhou, X., & Shu, H. (1999). Sublexical processing in reading Chinese: A development study. *Language and Cognitive Processes, 14*(5), 503-524. doi: 10.1080/016909699386176

Wu, Y. J., & Thierry, G. (2010). Chinese–English Bilinguals Reading English Hear Chinese. *The Journal of Neuroscience, 30*(22), 7646.

Wu, Y. J., & Thierry, G. (2010). Investigating bilingual processing: the neglected role of language processing contexts. *Frontiers in Psychology, 1*.

Wu, Y. J., & Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Frontiers in Psychology, 2*.

Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *Neuroimage, 59*(4), 3468-3473.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience, 170*(1), 223-231. doi: 10.1016/j.neuroscience.2010.06.077

Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2014). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics, 33*(0), 29-49. doi: 10.1016/j.jneuroling.2014.09.004

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhang, Q., & Zhu, X. (2011). The temporal and spatial features of segmental and suprasegmental encoding during implicit picture naming: An event-related potential study. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2011.09.040

Zhang, X. (1998). *Dialect MT: a case study between Cantonese and Mandarin.* Paper presented at the Proceedings of the 17th international conference on Computational linguistics-Volume 2.

Zhang, X., Samuel, A. G., & Liu, S. (2012). The perception and representation of segmental and prosodic Mandarin contrasts in native speakers of Cantonese. *Journal of Memory and Language, 66*(3), 438-457. doi: 10.1016/j.jml.2011.12.006

Zhao, J., Li, Q.-L., Wang, J.-J., Yang, Y., Deng, Y., & Bi, H.-Y. (2012). Neural basis of phonological processing in second language reading: An fMRI study of Chinese regularity effect. *Neuroimage, 60*(1), 419-425.

Zhou, X., & Marslen-Wilson, W. (1999a). The nature of sublexical processing in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 819-837.

Zhou, X., & Marslen-Wilson, W. (1999b). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language, 41*, 579-606. doi: 10.1006/jmla.1999.2663

Zhou, X., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1245-1265. doi: 10.1037/0278-7393.26.5.1245

Zinszer, B. D., Chen, P., Wu, H., Shu, H., & Li, P. (2014). Second language experience modulates neural specialization for first language lexical tones. *Journal of Neurolinguistics*.

Zinszer, B. D., Malt, B. C., Ameel, E., & Li, P. (2014). Native-likeness in second language lexical categorization reflects individual language history and linguistic community norms. *Language Sciences, 5*, 1203.

---

[i] Chapter 5 has been published: Wu, J., Chen, Y., Van Heuven, V. J., & Schiller, N. O. (2014). *Tonal variability in lexical access*. Language, Cognition and Neuroscience. doi: 10.1080/23273798.2014.915977.

Chapter 6 has been included in the proceeding of a conference: Wu, J., & Chen, Y. (2014). *Tonal variants in the bilingual mental lexicon.* Paper presented at the the Fourth International Symposium on Tonal Aspects of Languages (TAL 2014), Nijmegen.

Part of Chapter 3 has been presented as a poster in a conference: Wu, J., Chen, Y., & Schiller, N. O. (2015). *Aging and Age-Independent Effects of Cognitive and Sociolinguistic Backgrounds: on the Strength of Tonal Systematic Correspondence by Tonal Bilinguals*. Paper presented at the Bilingualism and Cognitive Aging, Groningen.

Chapters 2, 3, and 4 have been submitted to journals. Chapter 7 is in preparation for submission.

[ii] However, SC is not equivalent to Beijing Mandarin, because some of the morphological lexical variants (i.e. variants with ‘Erhua’ erization) and specific words from Beijing Mandarin were not introduced into SC in the standardization. Nowadays many young Beijing citizens also speak SC instead of real Beijing Mandarin or use accents closer to SC in formal situations.

[iii] Low-falling+Low sandhi pattern also exists in some speakers

[iv] The oldest speakers recorded by Qian (1997) still distinguished the vowel [ɛ] used in [tɕiɛ] ‘connect’ and the vowel [e] used in the other words of this minimal set. However, the younger speakers seem to use [ɛ] for all these words.

v In this book, the terms ‘tonal bilinguals’ and ‘tonal monolinguals’ are used to refer to SC-JM speakers from Jinan and SC speakers from Beijing. This terminology focuses on the speakers’ experience with tonal dialects. It does not rule out the possibility that both groups may speak other non-tonal languages with some proficiency, i.e. English, German, or French.

[vi] Tai and other tonal languages within China are in long-term close contact with local Chinese dialects (usually Southwestern Mandarin) and in the last 30 years also with Standard Chinese. There has been a long history of debate over the historical relationship between these languages and Chinese (Benedict, 1942; F.-k. Li, 1973; Matisoff, 2003). It is nevertheless clear that the dialects of these tonal languages within the borders of China share a considerable number of (etymologically) related words with Chinese (although it is still under debate whether these related words come from a common origin or early borrowing). Tonal bilingualism is a common phenomenon among the speakers of these languages in China.

2 Interlingual Two-to-One Mapping of Tonal Categories

Abstract

Both Standard Chinese (SC) high- and low-rising tones sound like the rising tone in Jinan Mandarin (JM). The role of this two-to-one interlingual tonal mapping was investigated in speech production, speech perception, and lexical access. Statistical modeling suggests that both SC rising tones overlap with the JM rising tone, with the high-rising SC tone having a better interlingual category-goodness. A semantic priming experiment focused on whether and how SC rising tones activate JM words in native bilinguals' mental lexicon and to what extent the interlingual category-goodness exerts its influence on speech comprehension. The bilinguals made auditory lexical decisions from a list including JM real words, JM non-words, and SC pseudo-words which sound like other JM real words. The SC pseudo-words (primes) were followed by the JM real words (targets). These targets were semantically related or unrelated to the JM real words hinted at by the SC primes. The SC pseudo-words ended with either high-rising or low-rising tones. The results showed that both high-rising and low-rising final SC pseudo-words were accepted as JM real words with a higher-than-chance probability. The activation of JM lexical items from SC tones was asymmetrical. SC pseudo-words ending with high-rising tones were more likely and more quickly recognized as real JM words, compared with their counterparts ending with low-rising tones. However, when accepted as JM real words, the final low-rising and final high-rising pseudo-words were equivalent in their semantic priming effects. These results support overlapping divisions of the tonal acoustic space according to the language mode in lexical access and some discreteness in the stage between lexical activation and semantic activation.

2.1 Introduction

Two different phonemes in one language matching to one and the same phoneme in the other language is a common phenomenon in bilingualism. For instance, Dutch and German learners have difficulty distinguishing English /æ/ and /ɛ/ because they only have /ɛ/ whose acoustic distribution primarily overlaps with, but is still different from the English /ɛ/ (Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997). Similarly, Japanese learners have difficulty distinguishing English /r/ and /l/ because they only have /ɾ/ which, while also apico-alveolar, is instead tapped (Best & Strange, 1992; Miyawaki et al., 1975). Such phenomena have been extensively investigated in second-language phoneme perception, and the related confusions in lexical access have also been studied.

In lexical decision, the minimal pairs, which are not contrastive in the native language become 'pseudo-homophones' (e.g. English locket vs. rocket for Japanese listeners) and prime each other like repetitions for the same word (Anne Cutler & Otake, 2004; Dufour, Nguyen, & Frauenfelder, 2007; Pallier, Colomé, & Sebastián-Gallés, 2001). 'Near-words' constructed by replacing a phoneme (e.g. English /t/) with its confusing phoneme (e.g. English /d/) are taken as words by listeners who have difficulty distinguishing the pair in their native language (e.g. Dutch /t/ and /d/ are neutralized in word-final position[1]); for the non-native listeners in repetition priming, such 'near-words' embedded in a context also prime the corresponding real word (e.g. 'half lied' primed 'flight') causing 'phantom word activation' (Broersma & Cutler, 2008).

The two-to-one interlingual mapping can be asymmetrical. For instance, the Japanese /ɾ/ is perceptually more similar to the English /l/ than the English /r/ (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Iverson et al., 2003). The phonetic asymmetry also affects how the phonetic representations activate lexical representations. In auditory picture-word identification, the picture of a 'locker' triggered more interference with the selection of the target 'rocket' than vice versa (A. Cutler, Weber, & Otake, 2006). A similar asymmetry was found between 'pencil' and 'panda' for Dutch-English bilinguals and Dutch learners of English (Escudero, Hayes-Harb, & Mitterer, 2008; Weber & Cutler, 2004). According to Cutler and Weber (2004, 2006), these results support two aspects of bilingual lexical access. First, the L2 acoustic input is captured by the native phonemic categories. Second, even though an inaccurate categorization may happen during the phonetic processing before lexical access, non-contrastive L2 phonemes in native perception are still stored with distinctive mental representations.

Similar two-to-one mapping patterns also abound in tonal languages and dialects. For instance, Jinan Mandarin (JM) has only one rising tone (JM T1) (Qian, 1997) but Standard Chinese(SC) has two rising tones, one high (SC T2) and one low (SC T3)(Moore & Jongman, 1997; Shen & Lin, 1991). This two-to-one tonal mapping is conditional. Specifically, this two-to-one cross-language tonal mapping is only valid in isolation and at prosodic word boundaries, where the SC high- and low-rising tones are contrastive but only one JM rising tone exists.

The two-to-one interlingual mapping of tones may nevertheless be different from the previously investigated segmental cases. First, lexical tone is a special type of phoneme in speech perception and lexical access. Although categorical perception of native tones is supported by recent neurological studies (Chandrasekaran, Krishnan, & Gandour, 2009; Xi, Zhang, Shu, Zhang, & Li, 2010), the native perception of Chinese tones in behavioral studies is 'quasi-categorical', neither as categorical as that of consonants nor as continuous as that of vowels (Hallé, Chang, & Best, 2004). Tonal information, compared with segmental information, is also retrieved later (Ye & Connine, 1999; Zhang & Damian, 2009; Zhang & Zhu, 2011) and involves different neuronal networks (Liang & van Heuven, 2004) in speech production. Considering lexical access, although lexical adaptation (McQueen, Cutler, & Norris, 2006; Mitterer, Chen, & Zhou, 2011) and constraining activation (Malins & Joanisse, 2010) seem to work similarly in tones and consonants, the overlap of SC tones alone induced no facilitatory priming effect in implicit priming, different from the classical effect observed by segmental primes (Chen, Chen, &

---

[1] According to (Warner, Jongman, Sereno, & Kemps, 2004), this neutralization of Dutch word-final /t/ and /d/ may be incomplete. Some sub-phonemic durational difference is maintained, which, although inconsistent in production, can still be noticed in perception.

Dell, 2002). In spite of these studies, the role of tone in perception and lexical access is little studied in bilinguals who use two tonal systems natively.

Second, aside from the particularity of tone compared with segmental phonemes, the two-to-one mapping between SC and JM rising tones seems different when regarding the overlapping of acoustic distributions. The Japanese /ɾ/ is a tap (Vance, 1987), while the English /l/ and /r/ are lateral and central approximants, respectively (Ladefoged, 2001). The north German /ε/ has lower F1 (less open) and higher F2 (more front) than the American English /ε/ (Strange, Bohn, Trent, & Nishi, 2004). Thus in both cases, the two-to-one mapping actually involves three similar but distinctive phones (i.e. the English /æ/, English /ε/[ε], and the German /ε/[ε̝]). However, is this also the case of SC-JM tonal mapping? The acoustic distribution of the JM rising tone seems to largely overlap with the acoustic distribution of both the SC high-rising and low-rising tones, although the mapping with the SC high-rising tone seems to be greater. This mapping asymmetry needs experimental verification before we make further investigations into their contribution to bilingual lexical access.

Third, different from the L2 learners in previous two-to-one mapping studies, the young SC-JM bilinguals are mostly simultaneous or early bilinguals with SC as the dominant language and received literacy education only in SC. These bilinguals thus have high proficiency in both dialects and have little difficulty in producing the correct rising tones according to the language mode and the lexical meaning.

How tonal bilinguals map tonal representations to lexical representations in speech perception and lexical access, however, needs more investigation. To investigate this two-to-one interlingual tonal mapping case, some other situations need to be controlled. On the one hand, the SC low-rising tone has different variants in non-final position. Before another low-rising tone, the SC low-rising and high-rising tones are nearly merged (Peng, 2000; Yuan & Chen, 2014) and not distinguishable in speech perception. The non-final tonal realizations both sound like the SC high-rising tone in isolation. For instance, 'land reform' and 'alter' are homophones in SC, due to the tone sandhi which happens on the first word /tʰu(low-rising) + kai(low-rising) -> tʰukai(high-rising+low-rising) = tʰu(high-rising) + kai(low-rising)/. Also, the rising part of the SC low-rising tone only appears in isolation at prosodic word boundaries (Garding, Kratochvil, Svantesson, & Zhang, 1986; Hallé et al., 2004). On the other hand, the JM rising tone realizes congruently higher in non-final positions and seems to only map to the SC high-rising tone. Thus, JM and SC rising tones have a one-to-one mapping (high-rising with high-rising) in non-final positions. To avoid the confusion from tone sandhi, SC stimuli with low-rising tone on the first syllable were avoided in the current study.

The acoustic overlapping between JM and SC rising tones gives rise to additional questions. How do the SC-JM bilinguals store the rising tones in their mental lexicon for auditory lexical access? More specifically, do they store the JM rising tone with a separate mental representation or do they store the JM rising tone integrated with the more similar SC high-rising tone? One possibility is that they store only two tonal representations, one high-rising and the other low-rising, but the categorization of the high-rising tone serves both JM and SC lexical access. If this hypothesis is true, we should expect the canonical realization of the SC low rising tone to fail matching any tonal representation in JM lexical access. The other possibility is that the bilinguals store three tonal representations. Then the theory needs to allow two tonal representations in the same mind to have overlapping acoustic distributions. This hypothesis also implies that the bilinguals can process the same tonal realization differently depending on the dialect in use. This is possible when considering the previous finding that Greek–English early sequential bilinguals gave different category-goodness ratings for the same physical stimuli depending on the language mode (Antoniou, Tyler, & Best, 2012). If this hypothesis is true, the canonical realizations of both the SC high-rising and low-rising tones could be taken as the JM rising tone in lexical access. Also, similar to the asymmetry found for both consonants and vowels (A. Cutler et al., 2006; Weber & Cutler, 2004), acoustic realizations of the SC high-rising tone, compared with that of the SC low-rising tone, should be better captured by the mental representation of the JM rising tone. Thus, with the segmental structure aligned, the SC pseudo-word with a high-rising tone should more likely be accepted as a real JM word. Also, the lexical decision of the SC high-rising pseudo-word should be faster, considering the asymmetry found in the bilingual lexical access involving asymmetrical two-to-one segmental mapping patterns (A. Cutler et al., 2006; Escudero et al., 2008; Weber & Cutler, 2004).

There is also a further question about the influence of the category-goodness. The canonical realization of the SC high-rising tone, compared with that of the SC low-rising tone, is a better exemplar of the JM rising tonal category. As discussed above, we expect this difference in category-goodness to influence the interlingual phantom activation of the JM lexical representations. However, to what extent does this category-goodness maintain its influence in speech comprehension? Researchers have a strong consensus that the phonological and semantic activation proceed in a largely parallel way in auditory speech comprehension. For instance, the ‘cohort model’ suggests that, as soon as the first phonemes of the target word are identified, the corresponding semantic information is activated to some extent (Grosjean, 1980; William Marslen-Wilson, 1984; W. D. Marslen-Wilson & Welsh, 1978). This parallel or cascading view of speech comprehension has been supported by both behavioral and neurophysiological evidence. For instance, shadowing can be performed before the utterance ends (W. Marslen-Wilson, 1973). Partial primes are enough to cause associative priming in sentential context (Zwitserlood, 1989). The semantic effect on the N400 component onsets before the offset of the eliciting word (Holcomb & Neville, 1991) and the latency of the N400 component is affected by whether the pseudo-word is different from the real word at the beginning (O'Rourke & Holcomb, 2002; Van Petten, Coulson, Rubin, Plante, & Parks, 1999). Comparing the timing of the N200 ERP components related to the monitoring of phonological and conceptual features in a Go/noGo paradigm, the former component precedes but largely overlaps with the later component (Rodriguez-Fornells, Schmitt, Kutas, & Münte, 2002). However, these findings did not distinguish lexical activation from semantic activation. If the tonal category-goodness influences lexical access, does this influence also reach the semantic level? If it does, SC pseudo-words with high-rising tone should also increase the semantic activation of the corresponding JM real word. Alternatively, if the influence of the category-goodness stops after lexical activation and does not spread to the semantic level, the SC pseudo-words with high-rising and low-rising tones should show no further difference in semantic activation, so long as they succeed in activating the JM lexical node. The strength of semantic activation can be tested with the semantic priming paradigm.

In the current study, three experiments tested the production, perception, lexical access, and semantic effects of the SC and JM rising tones. Experiment 1 compared the acoustic distribution of the two SC rising tones against the JM rising tone and verified the asymmetrical mapping in speech production. Experiment 2 tested SC tonal monolinguals' perception of the JM rising tone as SC tones and verified the asymmetrical mapping in SC speech perception. Experiment 3 tested SC-JM bilinguals' auditory lexical decision of SC pseudo-words (with high-rising and low-rising tones) as JM real words (with the JM rising tone), as well as their semantic priming effects on JM real words. The lexical decision of the SC pseudo-words as JM real words reflects the effect of interlingual tonal goodness in bilingual lexical access. We expect the SC high-rising pseudo-words to be more likely and quickly accepted as JM real words. By comparing the semantic priming effects of SC high-rising and low-rising pseudo-words we want to answer the question whether the influence of the tonal category-goodness in speech comprehension is maintained at the semantic level.

2.2 Experiment 1: acoustic distribution of JM and SC rising tones

Experiment 1 aimed at investigating the acoustic distributions of the JM rising tone and the two SC rising tones at the end of disyllabic word forms (JM words and SC pseudo-words).

2.2.1 Participants

Forty-two native JM tonal bilinguals from Jinan (16 male and 26 female, aged between 23 and 76 years, M = 40.29, SD = 17.04; seventeen SC dominant or balanced, twenty-five JM dominant) and 48 SC tonal monolinguals from Beijing (7 male and 41 female, aged between 19 and 30 years, M = 22.73, SD =2.95) participated in this experiment in exchange for payment.

2.2.2 Corpus preparation

The list of stimuli is composed of 16 final-rising disyllabic JM words and their corresponding SC pseudo-words, which either end with high-rising or low-rising SC tones (see the appendix for the full list).

The JM final-rising disyllabic words were selected from a 400-word corpus produced by 42JM speakers. The selection was controlled in that the targeted JM tonal patterns were produced by at least 88% of the JM speakers in our corpus and the JM words do not have false-friends in SC. In order to construct the corresponding SC pseudo-words, these disyllabic JM words also satisfy the criterion that their monosyllabic compositions have false-friends in SC and do not have frequent alternative pronunciation. Both JM and SC morphemes can be written with Chinese characters.

For each JM final-rising word, a pair of SC pseudo-words was constructed, which shares the same monosyllabic morpheme in the first syllable and ends with monosyllabic tonal minimal pairs carrying SC high and low-rising tones, respectively. For instance, a pair of SC disyllabic pseudo-words were constructed as ‘绳银’ /ʂəŋ in(high-rising+high-rising)/ and ‘绳引’ /ʂəŋ in(high-rising+low-rising)/,which are homophone candidates for the JM word ‘sound’ /ʂəŋ in([high-]rising+rising)/. The SC pseudo-words were then presented with Chinese characters and named by the SC tonal monolinguals.

For both the JM and SC participants, the printed stimuli were presented in a different random order for each speaker. After the speaker finished producing a word or pseudo-word, they pressed a key to see the next word. We used Praat (Boersma & Weenink, 2014) to extract pitch contours. Only pitch contours on the rhymes were extracted. A trained phonetician listened to each recording, looked at the spectrogram, and manually marked the rhyme of each syllable. Also, in this process, recordings with speech errors or recording errors were excluded from the corpus. Afterwards, the pitch contours were converted from hertz to semitones with 100Hz as the base and then transformed into z-scores based on the speakers’ means and standard deviations. This normalization removed the difference of pitch register across speakers, which is not the focus of the present study. The normalized pitch contours were then interpolated to 20 points per-syllable to remove the difference in duration.

2.2.3 Analysis and results

We built Generalized Additive Models (GAM) for the *Pitch* data, using the ‘mgcv’ package (S. Wood, 2006; S. N. Wood, 2011) in R (R_Core_Team, 2013). The data were stratified according to the JM tonal combinations (each together with their two corresponding SC tonal combinations) and fitted with separate models. In each model, a three- level factorial predictor *tone* [JM rising (JM t1), SC high-rising (SC t2), and SC low-rising (SC t3)] was included. Smooth functions were used to model non-linear functional relations between *Pitch* and the position of the point on the pitch contour (*pnt*). *Tone* (the JM or SC tone carried by the ending syllable) was included in both the fixed linear predictors and fixed smoothes. The candidates for random predictors were *itemID*, *set ID* (each JM real word and its similar SC final high-rising and low-rising pseudo-words form a set), and *Speaker*. Since *item ID* is nested under *set ID* and predictable due to the combination of *set ID* and *tone*, we built models which would otherwise be identical, including *item ID* and *set ID* together, *set ID* alone, or *item ID* alone in the random terms. The structure of the final model was decided by model comparison based on the Akaike Information Criterion (AIC) likelihood values (Sakamoto & Ishiguro, 1986). The models with the *item ID* alone turned out to be the best. Thus, *item ID* but not *set ID* was included in the random terms. After the structure of the model was decided, autocorrelation values were calculated based on the order of data points in the pitch contour and the greatest value was included as the AR1 correlation parameter to build the corresponding AR1 error model (S. Wood, 2006; S. N. Wood, 2011) but when the AR1 error model does not improve the original model, the original model was reported.

The final models for different SC tonal combinations shared the same structure. *Tone* was included as the factorial predictor; a thin-plate regression spline smooth was included to model the interaction of *pnt* and *tone*; the smoothes of the by-*Speaker* random slope of *pnt* and the smooth of by-*word ID* random slope of *pnt* were included as the random predictors.

The fitted models accounted for 76.1% of the variance in the data of the JM and SC rising + (high/low) rising tonal combination, 86.7% of the variance in the data of the JM and SC high-level + (high/low) rising tonal combination, 76.4% of the variance in the data of the JM and SC falling + (high/low) rising tonal combination. The coefficients for the parametric predictors are shown in Table 1. The number of degrees of freedom in the smooth terms and the associated F-statistics are shown in Table2. The fitted pitch contours are shown in Figure1.

As shown by the scattered contour plots in Figure 1 (van Rij, Wieling, Baayen, & van Rijn, 2015), the JM rising tone (JM T1) overlaps with both SC rising tones in the acoustic distribution, with greater overlap with the SC high-rising tone (SC T2). As shown by the estimated smoothes in Figure 1, SC final low-rising (SC T3) pseudo-words carry lower pitch contours on the second syllable than SC final high-rising (SC T2) pseudo-words and the contours of JM final-rising (JM T1) words lie in between the two SC rising tones. With the rising (JM T1 & SC T2) and high-level (JM T3 & SC T1) tones on the first syllable, although the JM tones seem to cover a slightly smaller and lower pitch register, the cross-dialect mapping is very close. However, the pitch register of JM low-falling (JM T4) on the first syllable is much lower than that of the SC falling (SC T4). Hence, it is mainly the shape of the pitch contour instead of the pitch register that is similar between the JM and SC falling + rising combinations (JM T4 + T1& SC T4 + T2/T3). Also, the shape of this JM tonal combination is more similar to the final high-rising SC counterpart than to the other tonal combinations.

Table 1. Coefficients for the linear predictors in the generalized additive model fitted to *Pitch* of JM & SC production data (** $p < 0.01$, *** $p < 0.001$).

	Predictors	Estimate	Std. Error	t value
JM T1 + T1 & SC T3 + T2/T3	(Intercept)	-0.2896	0.1545	-1.875 (n.s.)
rising + (high/low) rising:	toneSCt2	0.2642	0.2256	1.171 (n.s.)
	toneSCt3	-0.6342	0.2262	-2.804**
JM T3 + T1 & SC T1 + T2/3	(Intercept)	-0.02666	0.09758	-0.273 (n.s.)
high-level + (high/low) rising:	toneSCt2	0.41335	0.13908	2.972**
	toneSCt3	-0.15217	0.13913	-1.094 (n.s.)
JM T4 + T1 & SC T4 + T2/3	(Intercept)	-0.61862	0.08596	-7.197***
low-falling + (high/low) rising:	toneSCt2	0.51235	0.12191	4.203***
	toneSCt3	-0.09893	0.12202	-0.811(n.s.)

Table 2. Coefficients for the smooth terms in the generalized additive model fitted to *Pitch* of JM & SC production-identification data (*** $p < 0.001$).

	Smooth terms	edf	Ref.df	F
JM T1 + T1 & SC T3 + T2/T3 rising + (high/low) rising:	s(pnt):toneJMt1	16.41	16.46	680.99***
	s(pnt):toneSCt2	15.97	16.41	29.67***
	s(pnt):toneSCt3	16.70	16.83	141.81***
	s(pnt, Speaker)	708.31	763.00	353.04***
	s(pnt, itemID)	81.94	105.00	806.76***
JM T3 + T1 & SC T1 + T2/3 high-level + (high/low) rising:	s(pnt):toneJMt1	16.86	16.88	694.26***
	s(pnt):toneSCt2	16.52	16.87	50.37***
	s(pnt):toneSCt3	16.84	16.96	181.25***
	s(pnt, Speaker)	692.59	763.00	295.54***
	s(pnt, itemID)	104.15	132.00	436.13***
JM T4 + T1 & SC T4 + T2/3 low-falling + (high/low) rising:	s(pnt):toneJMt1	16.8	16.83	403.6***
	s(pnt):toneSCt2	15.2	16.32	44.64***
	s(pnt):toneSCt3	16.91	16.98	281.07***
	s(pnt,Speaker)	709.57	763.00	490.91***
	s(pnt, itemID)	153.69	186.00	67.48***

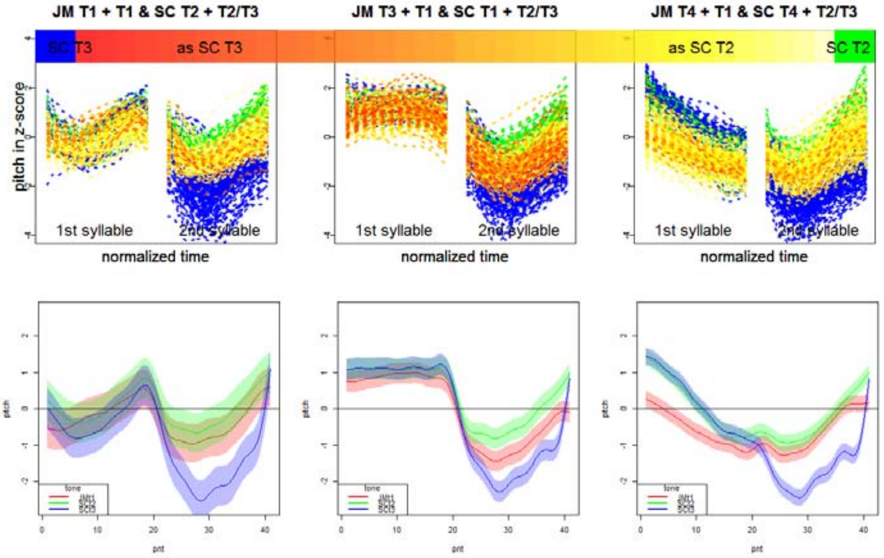


Figure 1. *Scattered contour plots (upper panel) and estimated smoothes (lower panel, random effect removed) for the JM final-rising words [JM T1, scaled from red (contour identified as SC T2) to yellow (contour identified as SC T3)], SC final-high-rising pseudo-words (SC T2, green), and S final-low-rising pseudo-words (SC T3, blue).*

2.2.4 Discussion

Experiment 1 verified that the acoustic distribution of the JM rising tone largely overlaps with both the SC high-rising and low-rising tones at word final position, with the overlap with the SC high-rising tone larger than the overlap with the SC low-rising tone. Despite the distributional overlapping, the GAM modeling showed that the three rising tones have different distributional centers, with the distributional center of the JM rising tone lying between the SC high-rising and low-rising tones.

2.3 Experiment 2: identification of JM words as SC pseudo-words

In Experiment 2, we verified the claim that the acoustic realizations of the JM rising tone can match both SC high-rising and low-rising tones in SC speech perception and investigated the relation between interlingual identification and the shape of pitch contours.

2.3.1 Participants

The forty-eight SC tonal monolinguals who participated in Experiment 1 also participated in Experiment 2.

2.3.2 Design and Stimuli

The 16 final rising disyllabic JM words (as shown in Appendix) produced by the 42 JM speakers in Experiment 1, with production errors and non-dominant variants excluded, were used as the stimuli in Experiment 2. Additionally, another SC-JM tonal bilingual who is highly proficient in both dialects produced the corresponding SC pseudo-words for each of the 16 JM words, which served as the training stimuli.

2.3.3 Procedure

The SC tonal monolingual participants performed a tonal identification task upon JM auditory stimuli using the E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) in a quiet room. With the JM words presented binaurally in random order, the participants judged which of the two corresponding SC pseudo-words printed on the screen was heard. The participants had 5,000 ms to make the judgment and the following stimulus appeared 1,000 ms after the response. Thirty-two training trials with real SC pseudo-words were tested before the crucial trials with feedback. It was verified that the participants were able to identify the training SC pseudo-words with high accuracy (ACC) [grand mean ACC = 0.97 (SD = 0.17), by-item ACC ranged from 0.74 to 1, by-participant ACC ranged from 0.91 to 1].

2.3.4 Analysis

We built Generalized Additive Models (GAM) for the *Pitch* data, using the ‘mgcv’ package (S. Wood, 2006; S. N. Wood, 2011) in R (D. M. Bates, 2010). Separate models were built for different JM tonal combinations and smooth functions were used to model non-linear functional relations between the predictors and the measurement *Pitch*. The SC tonal monolinguals’ choice for each stimuli (*Choice*), and position of the point in the pitch contour (*pnt*) were included as the fixed predictors, and *wordID*, *Speaker*, and *Participants* were included as the candidates for random predictors. The structure of the final model was decided by model comparison based on the AIC likelihood values (Sakamoto & Ishiguro, 1986). *Participant* did not reach significance or improve the model and was removed. Similar to the procedure used in Experiment 1, autocorrelation values were calculated and included as the AR1 correlation parameter to build the corresponding AR1 error model (S. Wood, 2006; S. N. Wood, 2011) but when the AR1 error model does not improve the original model, the original model was reported.

The final models for different SC tonal combinations shared the same structure. *Choice* was included as the factorial predictor; a thin-plate regression spline smooth was included to model the interaction of *pnt* and *Choice*. The smooth of the by-*Speaker* random slope of *pnt* and the smooth of the by-*word ID* random slope of *pnt* were included as the random predictors.

2.3.5 Results

The fitted models accounted for 75% of the variance in the data of the JM rising + rising tonal combination, 86.5% of the variance in the data of the JM high-level + rising tonal combination, 74% of the variance in the data of the JM low-falling + rising tonal combination. The coefficients for the parametric predictors are shown in Table 3. The numbers of degrees of freedom in the smooth terms and the associated F-statistics are shown in Table 4. The fitted pitch contours of the JM stimuli identified as final high-rising and low-rising SC pseudo-words and the corresponding difference contours are shown in Figure 2, Figure 3, and Figure 4. As shown in these figures, the differences of pitch contours between the stimuli identified as final high-rising and final low-rising (SC T2 vs. T3) SC pseudo-words are consistent. Stimuli with relatively lower pitch contours on the second syllable are more likely to be identified as carrying the SC low rising tone (SC T3) on that syllable. Also, the stimuli with higher pitch contours on the first syllable are more likely to be identified as carrying the SC low rising tone on the second syllable.

GAM models similar to those built in Experiment 1 yielded the plots in the lower panels of Figures 2, 3, and 4. As shown in these plots and the scattered contours in Figure 1, the SC final low-rising (SC T3) pseudo-words carry lower pitch contours on the second syllable than the SC final high-rising (SC T2) pseudo-words. The shape of the difference curves is consistent with the difference curves from the perceptual data of JM real words. However, the SC final low-rising pseudo-words do not always carry higher pitch contours on the first syllable for compensation. Also, it seems that the SC monolinguals’ identification of JM tones as SC tones is mostly based on the shape of the pitch contour. As was also shown in the results of Experiment 1 (Figure 1), the pitch contour of JM low-falling (JM T4) is much lower than that of the SC falling (SC T4). Although the general pitch register is different, the relative pitch register of JM low-falling (JM T4), compared with that of JM rising (JM T1), makes the shape of pitch contour of the JM low-falling + rising (JM T4 + T1) combination more similar to that of the SC falling + high-rising combination. As shown in Figure 4, the difference curve between the two identified subcategories is more similar to that of the other JM tonal combinations than the corresponding SC difference curve. Also, a majority of the JM T4 + T1 stimuli were identified as SC final high-rising pseudo-words.

Table 3. Coefficients for the linear predictors in the generalized additive model fitted to *Pitch* of JM production-identification data (** p < 0.01, *** p < 0.001).

	Predictors	Estimate	Std. Error	t value
JM T1 + T1	(Intercept)	0.120992	0.288234	0.42(ns.)
rising + rising:	Choice: as T3	-0.030936	0.005596	-5.529***
JM T3 + T1	(Intercept)	0.17975	0.18029	0.997 (ns.)
high-level + rising:	Choice: as T3	-0.02345	0.00488	-4.805 ***
JM T4 + T1	(Intercept)	-0.453184	0.133084	-3.405***
low-falling + rising:	Choice: as T3	-0.041467	0.005352	-7.748***

Table 4. Coefficients for the smooth terms in the generalized additive model fitted to *Pitch* of JM production data (*** p < 0.001).

	Smooth terms	edf	Ref.df	F
JM T1 + T1	s(pnt):Choice: as SC T2	16.04	16.05	1082.00***
rising + rising:	s(pnt):Choice: as SC T3	16.15	16.17	269.80***
	s(pnt, Speaker)	366.20	377.00	299.30***
	s(pnt, Word ID)	30.69	35.00	829.00***
JM T3 + T1	s(pnt):Choice: as SC T2	16.61	16.62	1329.70***
high-level + rising:	s(pnt):Choice: as SC T3	16.62	16.64	551.20***
	s(pnt, Speaker)	365.82	377.00	237.80***
	s(pnt, Word ID)	38.54	44.00	640.90***
JM T4 + T1	s(pnt):Choice: as SC T2	16.62	16.64	1248.50***
low-falling + rising:	s(pnt):Choice: as SC T3	16.57	16.62	144.10***
	s(pnt, Speaker)	366.56	377.00	456.50***
	s(pnt, Word ID)	56.46	62.00	923.30***

2.3.6 Discussion

Experiment 2 verified that the JM rising tone matches both the SC high-rising and low-rising tones in interlingual speech perception by naïve SC listeners. It is also found that the interlingual tonal identification is biased by the pitch height in the specific rendition of the word: when JM final rising words carry relatively lower ending pitch contours, they were more likely to be identified as the SC final low-rising pseudo-words. The pitch height in the first syllable also slightly affects the interlingual tonal perception, in that the higher the previous pitch is, the more likely the word is to be identified as the SC final low-rising pseudo-words.

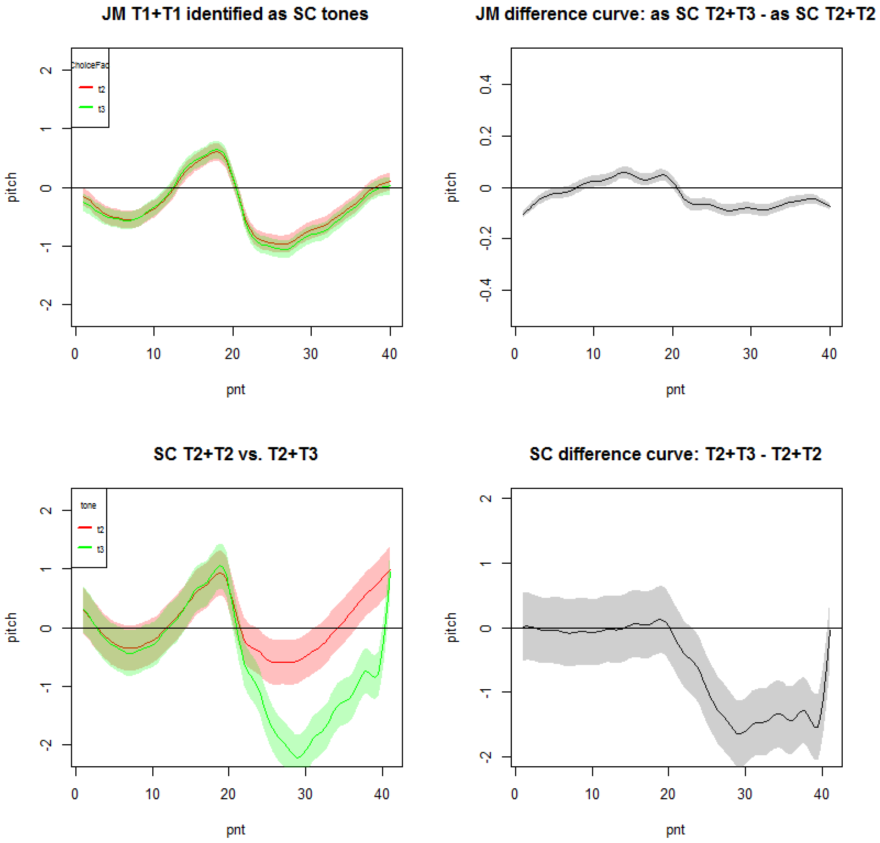


Figure 2. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T1+T1) identified as SC final-high-rising (SC T2+T2) and final-low-rising (SC T2+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T2+T2) and final-low-rising (SC T2+T3) pseudo-words.*

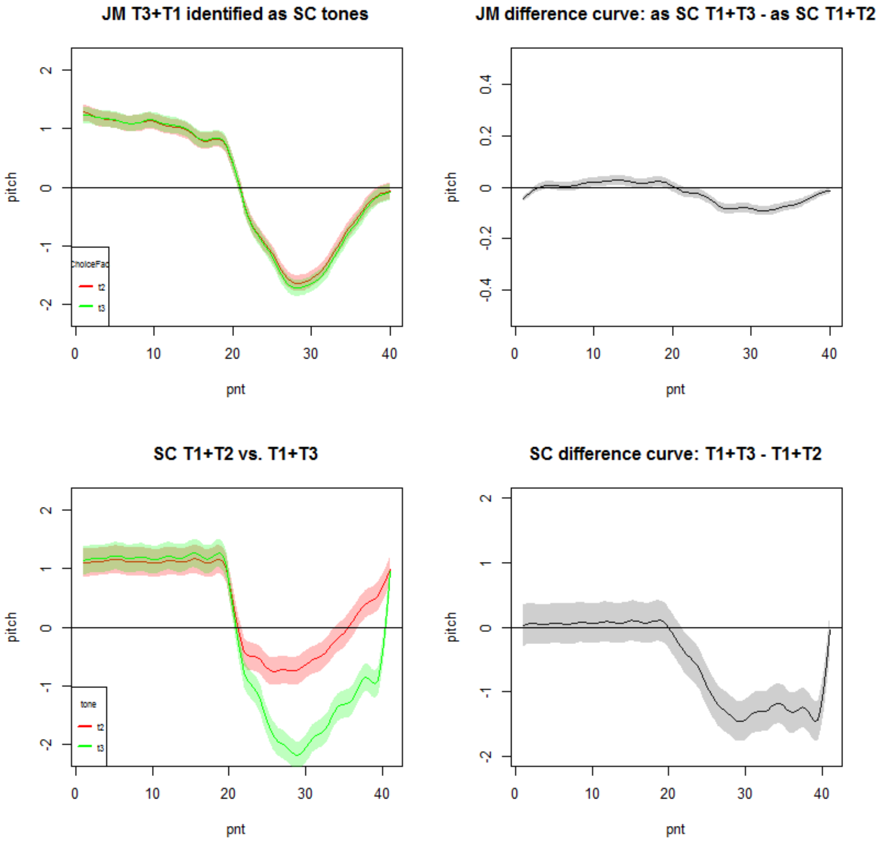


Figure 3. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T3+T1) identified as SC final-high-rising (SC T1+T2) and final-low-rising (SC T1+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T1+T2) and final-low-rising (SC T1+T3) pseudo-words.*

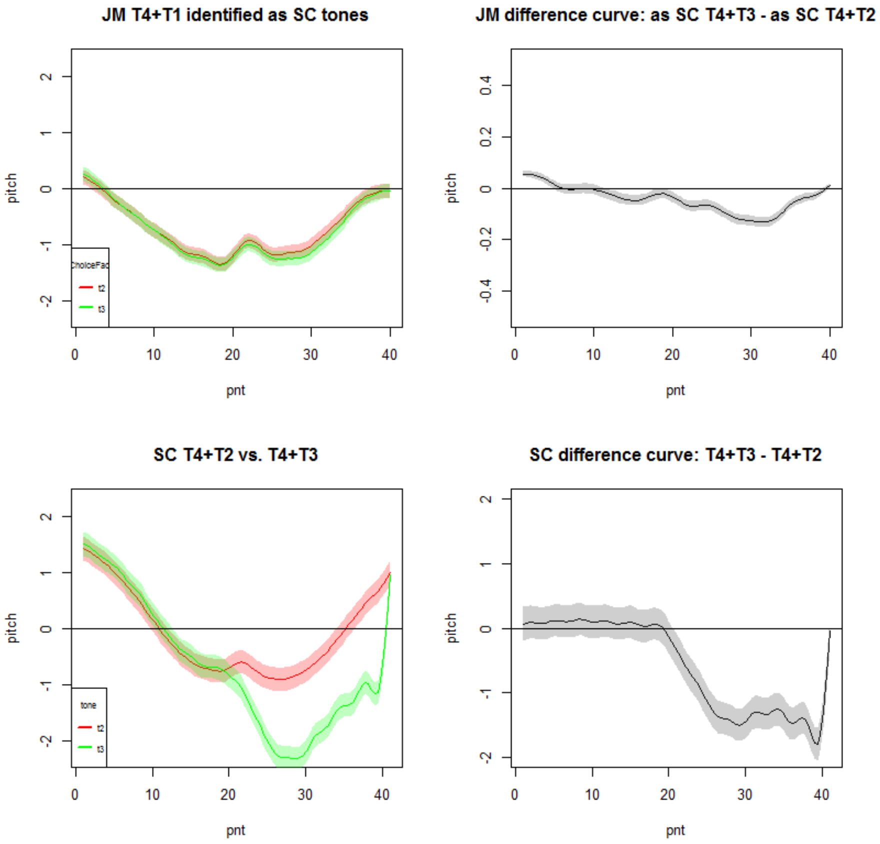


Figure 4. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T4+T1) identified as SC final-high-rising (SC T4+T2) and final-low-rising (SC T4+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T4+T2) and final-low-rising (SC T4+T3) pseudo-words.*

2.4 Experiment 3: bilingual semantic priming

Experiment 3 used the lexical decision task and semantic priming paradigm to investigate whether the two SC rising tones can both activate SC-JM tonal bilinguals' JM lexical representations, whether the mapping goodness of the two SC tones to the JM tonal categories affects the speed of JM lexical activation, and whether the mapping goodness also affects the semantic activation of the JM words.

2.4.1 Participants

Fifty-five native SC-JM tonal bilinguals from Jinan (15 male and 40 female, aged between 19 and 36 years, M = 23, SD = 3.85; 45 SC-dominant or balanced, 10 JM-dominant) participated in this experiment in exchange for payment. All participants were right-handed, received their literacy educations in SC, and learned some English at school. Four participants from the Jinan group and 15 participants from the Beijing group also had some knowledge of other non-tonal foreign languages, such as French and German.

2.4.2 Design and Stimuli

The present study adopted the semantic priming paradigm. Pairs of SC disyllabic pseudo-words were constructed with non-sense combinations of Chinese characters as primes, with one member ending with a high-rising tone and the other member ending with a low-rising tone. Each pair of pseudo-words was constructed so that their SC pronunciations may be mapped onto the same JM real words, which carry final rising tones. The matching JM words were selected from a 400-word corpus produced by 42 JM speakers (as used in Experiments 1 and 2). The two-to-one interlingual tonal mapping seems to cause no problem for the younger generations of SC-JM bilinguals.

To select semantically related targets, we combined the procedures used in earlier studies (Sumner & Samuel, 2005; Thierry & Wu, 2007). First a printed list of the primes was presented to a different group of 16 native SC speakers who were instructed to write down a related word for each item. One or two related targets were chosen for further selection based on the number of participants who wrote the given target as response. Then we formed word-pairs by crossing these potential targets with the primes and the semantic relatedness of these word pairs were rated on a scale from 1 to 5 by a group of 20 native SC speakers, including some of the above-mentioned native SC speakers and some new raters. One semantically related target and one semantically unrelated target were accordingly selected for each prime. For each JM real word, we used its related SC high-rising pseudo-word, its related SC low-rising pseudo-word, its semantically related target, and its semantically unrelated target as stimuli, as shown in the appendix together with the semantic relatedness data. The stimuli also included 32 JM non-words. A male native bilingual who is highly proficient in both dialects then produced these words and non-words accordingly in JM and SC (also a trained phonetician with Putonghua Proficiency Test Certificates- Level1B).

The design was a four-by-four Latin square design. Test sets were split into four lists, participants were also split into four groups, and the combination of prime conditions (final high-rising or low-rising) and target conditions (semantically related or semantically unrelated) were counterbalanced across the participants and test lists, so that each participant experienced every condition in the same number of trials and heard one prime and one target from each set.

2.4.3 Procedure

Participants were tested individually in a quiet room using the E-Prime software (Schneider et al., 2002). They were told that they would hear a series of sound sequences and that they were required to decide whether or not each of these sound sequences was a real word. Each item was played binaurally through headphones, with instructions on the screen. A new trial started 1,000 ms after the participant responded to the previous trial, or 1,000 ms after the response time exceeded 5 s. Each target was played directly after its corresponding prime with 1,000-ms ISI. The prime-target pairs were separated with non-word fillers and no prime-target pair was directly followed by another prime-target pair. In this way, we tried to control the potential phonological priming between the primes. The crucial test was preceded by a practice block including 5 words and 5 non-words.

2.4.4 Analysis and results

Analysis 1: Word-identification rates and reaction times to the primes. A *word-identification rate* was defined as the probability that a SC pseudo-word was identified by the SC-JM bilinguals as a real JM word. The average word-identification rates of both final high-rising and low-rising primes (78.18% and 67.95%, respectively) are above 50%. The identification data were collapsed into percentage scores indicating by-participant and by-item word-identification rate. Both by-participant and by-item t-test showed that the high-rising SC pseudo-words were more likely be identified as real JM words, by-participant t (54) = 3.38, $p < 0.01$, by-item t (15) = 2.98, $p < 0.01$.

Only the reaction times to the SC pseudo-words, which were accepted as real JM words, were taken into consideration. The reaction time data were also collapsed by participant and by item. Both by-participant and by-item t-test showed that the bilinguals responded faster to the high-rising SC pseudo-words than to the low-rising ones, by-participant t (54) = -5.49, $p < 0.001$, by-item t (15) = -5.20, $p < 0.001$.

These results indicate that both final high-rising and low-rising SC pseudo-words can be accepted as JM real words, although the final high-rising pseudo-words were more quickly and more likely to be accepted as real JM words.

Analysis 2: Accuracies and reaction times to the targets. The responses to the targets were influenced by more factors than the primes. We built Linear Mixed Effect (LME) models for the accuracy (ACC) and reaction time (RT) data (D. Bates, Maechler, Bolker, & Walker, 2013; R_Core_Team, 2013), including *Prime Condition* (final high-rising/final low-rising), *Target Condition* (related/unrelated), and their interactions as the fixed predictors. The candidates for the random terms included by-participant and by-target random intercepts and by-prime random intercepts nested under the related JM real word, as well as possible random slopes. The structure of the random terms in the models reported here was selected via model comparison based on likelihood ratio tests.

Logistic Linear Mixed Effect (LME) models were built for the binominal accuracy data. The selected random predictors were by-participant and by-target random intercepts, $X^2_{1|\text{participant}} = 7.44$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{target}} = 2.95$, $p_{1|\text{target}} < 0.1$. Parametric bootstraps (Singmann, 2014) showed that the main effect of *Target Condition* was significant, $F = 9.67$, $p < 0.01$. However, the main effect of *Prime Condition*, $F = 0.06$, n.s., and the interaction of *Prime Condition* and *Target Condition*, $F = 0.05$, n.s, were insignificant. Compared with the semantically unrelated targets, target $\text{ACC}_{\text{high-rising-prime}} = 96.4\%$, target $\text{ACC}_{\text{low-rising-prime}} = 96.4\%$, the semantically related targets showed significantly higher accuracy rates, target $\text{ACC}_{\text{high-rising prime}} = 99.5\%$, target $\text{ACC}_{\text{low-rising prime}} = 99.1\%$.

Linear Mixed Effect (LME) models were built for the reaction times to the targets. Only correct responses to the targets were considered and the reaction times were log-transformed to improve the distribution of the data. In the following analysis, a model was fit with all the data points, and then a model criticism removed the data points with standardized residuals exceeding 2.5 standard deviation units from the data set (less than 2.5% of the data) and refitted the model with the trimmed data set. We report the model statistics from the trimmed models, with Satterthwaite approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2013).

The first LME models were built for the responses collected after the corresponding primes were identified as real JM words. The selected random predictors were by-participant, by-related-JM real words, and by-target random intercepts, $X^2_{1|\text{participant}} = 124.64$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{related-JM real-word}} = 6.84$, $p_{1|\text{related-JM real-word}} < 0.01$, $X^2_{1|\text{target}} = 47.99$, $p_{1|\text{target}} < 0.001$. As shown in Figure 5, the main effect of *Target Condition* was significant, F ($df = 543.73$) $= 59.04$, $p < 0.001$. However, the main effect of *Prime Condition*, F ($df = 537.36$) $= 0.61$, n.s., and the interaction of *Prime Condition* and *Target Condition* was insignificant, F ($df = 532.63$) $= 0.00$, n.s. Compared to the semantically unrelated targets, the semantically related targets were processed faster.

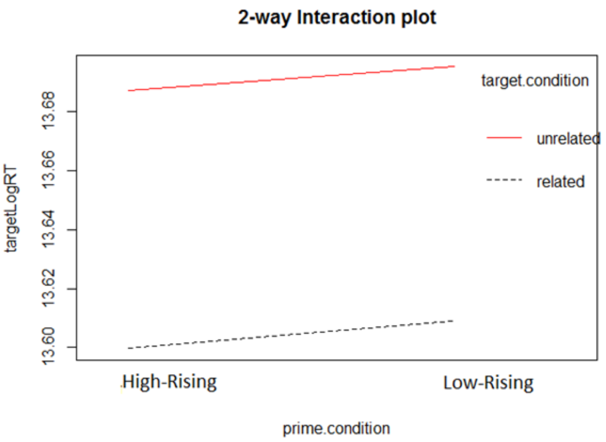


Figure 5. *Interaction plot of Prime Condition (columns) and Target Condition (lines) for Experiment 3 – Analysis 2.*

The second LME models were built for the responses collected after the corresponding primes were rejected as non-words. The selected random predictors were by-participant and by-target random intercepts, $X^2_{1|\text{participant}} = 49.75$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{target}} = 61.10$, $p_{1|\text{target}} < 0.001$. However, none-of the fixed predictors was significant, $F_{\textit{Prime-Condition}}$ (df=174.13) = 0.29, n.s., $F_{\textit{Target-Condition}}$ (df = 190.10) = 0.98, n.s., $F_{\textit{Prime}\times\textit{Target}}$ (df = 170.37) = 1.54, n.s. The reaction time was slightly reduced for the semantically related targets primed by final high-rising pseudo-words, although the difference was insignificant.

In sum, both final high-rising and final low-rising SC pseudo-words improved the accuracy and reduced the reaction times of the targets which are semantically related to the corresponding JM real words of the primes. The semantic priming effect was only salient when the SC pseudo-word primes were accepted as real JM words by the bilinguals.

2.5 Discussion and conclusion

2.5.1 Main results and interpretation

Experiments 1 and 2 together clarified the acoustic distribution and perceptual status of the SC and JM rising tones. Experiment 1 verified the impression that the acoustic distribution of the JM rising tone largely overlaps with both the SC high-rising and low-rising tones. As shown in Figure 1, the overlap with the SC high-rising tone is larger than the overlap with the SC low-rising tone. The GAM modeling also showed that the distributional center of the JM rising tone lies between the SC high-rising and low-rising tones. Experiment 2 verified that SC native monolinguals can perceive the JM rising tone as either the SC high-rising or low-rising tone. The interlingual tonal identification is largely based on how high the pitch is in the specific rendition. JM final rising words were more likely to be identified as the SC final low-rising pseudo-words when the syllable carried a relatively lower pitch contour. We also found that the asymmetry is more pronounced in the interlingual perception of the JM low-falling + rising combination as SC falling + high-/low-rising combination. This is probably because the JM low-falling is lower than the SC falling tone, which not only caused the following JM rising to be perceived as relatively higher, but also caused the whole pitch contour to match the SC high-rising even better.

Experiment 3 first verified that both the SC high-rising and low-rising tones can be accepted as the JM rising tone in JM lexical access and activate JM lexical nodes. The final high-rising SC pseudo-words were more quickly and more likely to be accepted as real JM words. Thus, the asymmetry also persists in interlingual lexical access. However, the asymmetry is not kept in the corresponding semantic activation. As long as the prime was accepted as a JM real word, it primed the semantically related target but whether the prime carried high-rising or low-rising tone makes no difference.

2.5.2 Theoretical implication

The interlingual perception of the tonal two-to-one mapping is consistent with the previous findings based on vowels and consonants (Aoyama et al., 2004; Best & Strange, 1992; Bohn & Flege, 1990; Flege et al., 1997; Iverson et al., 2003). The asymmetry in two-to-one interlingual mapping not only exists in vowels (Bohn & Flege, 1990; Flege et al., 1997) and consonants (Aoyama et al., 2004; Iverson et al., 2003), but also exists in tones. The asymmetry in tonal mapping has acoustic basis and perceptual effects, just as in segmental mapping.

The pitch register of the previous syllable serves as a reference in interlingual tonal identification. This finding is consistent with the previous findings. The same physical stimuli can be perceived as higher when the previous stimuli are low in pitch. This applies to both acoustic pitch perception and monolingual tonal perception (Fox & Qi, 1990; Leather, 1983; Lin & Wang, 1984; Moore & Jongman, 1997; Wong & Diehl, 2003; Wu, 2011). Obviously this also applies to interlingual tonal perception.

Previous studies on vowels and consonants support that the L2 acoustic input is captured by the native phonemic categories in lexical access (A. Cutler et al., 2006; Weber & Cutler, 2004). Also, it has been shown that Greek–English early sequential bilinguals gave different consonantal category-goodness ratings for the same physical stimuli depending on the language mode (Antoniou et al., 2012). However, there is limited evidence in support of overlapping divisions of the tonal acoustic space according to the language mode in lexical access. The present study first verified that the acoustic distribution of the JM rising tone overlaps with that of both SC rising tones and then found that the canonical realizations of the SC high-rising and low-rising tones could both be accepted as the JM rising tone in lexical access. Thus, two physical stimuli belonging to different tonal categories in one dialect can be captured by the same tonal category in the other dialect. The SC-JM early bilinguals do not store the JM rising tone integrated with the more similar SC high-rising tone. Instead, our finding supports the claim that two tonal representations belonging to different tonal systems in the bilingual mind can be associated with highly overlapping acoustic distributions.

Effects of consonantal and vowel asymmetry have been shown in interlingual lexical access in previous studies (Escudero et al., 2008; Weber & Cutler, 2004, 2006). The current study showed that the tonal asymmetry also affects how the tonal representations activate lexical representations. Since the SC high-rising tone overlaps more with the JM rising tone in acoustic distribution and its distributional center is closer to that of the JM rising tone, it is not surprising to find that the final high-rising SC pseudo-words are more likely to activate the corresponding JM real words, and do so faster. The current study supported that the asymmetric interlingual lexical access also involves tonal categories.

Using a semantic priming paradigm, the current study further distinguished the effect of phonetic asymmetry in interlingual lexical activation and semantic activation. When priming SC pseudo-words with JM real words, the semantic priming effect is symmetric and not affected by the interlingual category-goodness. Although it is strongly supported in the previous studies that the phonological and semantic activations happen in parallel in speech comprehension (Grosjean, 1980; William Marslen-Wilson, 1984; W. D. Marslen-Wilson & Welsh, 1978; O'Rourke & Holcomb, 2002; Rodriguez-Fornells et al., 2002; Van Petten et al., 1999; Zwitserlood, 1989), the influence of interlingual category-goodness is kept in lexical access but lost in semantic activation. This finding suggests that there is some discreteness in the step between lexical activation and semantic activation.

To sum up, the current study investigated the two Standard Chinese rising tones and their corresponding Jinan Mandarin rising tones in speech production, perception, and lexical access. The three rising tones form the two-to-one interlingual mapping pattern. The JM rising tone overlaps with both SC rising tones in speech production and speech perception and the interlingual mapping is asymmetric. This asymmetry affects interlingual lexical access by SC-JM bilinguals and supports the theoretical claim that the tonal acoustic space can be divided in an overlapping way in bilingual lexical access. The acoustic input is captured by the native tonal categories depending on the language mode. However, the asymmetric mapping does not affect the semantic activation after lexical access, which suggests some discreteness in the step between lexical activation and semantic activation.

References

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics, 40*(4), 582-594. doi: 10.1016/j.wocn.2012.05.005

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r. *Journal of Phonetics, 32*(2), 233-250.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Bates, D. M. (2010). lme4: Mixed-effects modeling with R. *URL http://lme4. r-forge. r-project. org/book.*

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] Version 5.3.82 retrieved 26 July 2014 from http://www.praat.org/.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(03), 303-328. doi: 10.1017/S0142716400008912

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System, 36*(1), 22-34. doi: 10.1016/j.system.2007.11.003

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*(1), 1-9. doi: 10.1016/j.bandl.2008.02.001

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751-781. doi: 10.1006/jmla.2001.2825

Cutler, A., & Otake, T. (2004). Pseudo-homophony in non-native listening. *The Journal of the Acoustical Society of America, 115*(5), 2392-2392. doi: 10.1121/1.4780547

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284. doi: 10.1016/j.wocn.2005.06.002

Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America, 121*(4). doi: 10.1121/1.2710742

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics, 36*(2), 345-360. doi: 10.1016/j.wocn.2007.11.002

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470. doi: 10.1006/jpho.1997.0052

Fox, R., & Qi, Y. Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics, 18*(2), 261-284.

Garding, E., Kratochvil, P., Svantesson, J. O., & Zhang, J. (1986). Tone 4 and Tone 3 discrimination in modern Standard Chinese. *Language and speech, 29*(3), 281. doi: 10.1177/002383098602900307

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*(4), 267-283. doi: 10.3758/BF03204386

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*(3), 395-421. doi: 10.1016/S0095-4470(03)00016-0

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286-300. doi: 10.3758/BF03332082

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57. doi: 10.1016/S0010-0277(02)00198-1

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf

Ladefoged, P. (2001). *A Course in Phonetics* (4th ed.): Heinle & Heinle Pub.

Leather, J. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics, 11*(4), 373-382.

Liang, J., & van Heuven, V. J. (2004). Evidence for separate tonal and segmental tiers in the lexical specification of words: A case study of a brain-damaged Chinese speaker. *Brain and Language, 91*(3), 282-293. doi: 10.1016/j.bandl.2004.03.006

Lin, T., & Wang, W. (1984). Shengdiao ganzhi wenti [Perception of tones]. *Zhongguo Yuyan Xuebao [Bulletin of Chinese Linguistics], 2*, 59-69.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62*(4), 407-420. doi: 10.1016/j.jml.2010.02.004

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*(5417), 522-523. doi: 10.1038/244522a0

Marslen-Wilson, W. (1984). Function and process in spoken word recognition-A tutorial review. *Attention and performance, 10*, 125-150.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*(1), 29-63. doi: 10.1016/0010-0285(78)90018-X

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126. doi: 10.1207/s15516709cog0000_79

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological Abstraction in Processing Lexical‐Tone Variation: Evidence From a Learning Paradigm. *Cognitive Science, 35*(1), 184-197. doi: 10.1111/j.1551-6709.2010.01140.x

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*(5), 331-340. doi: 10.1111/j.1551-6709.2010.01140.x

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America, 102*, 1864. doi: 10.1121/1.420092

O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological psychology, 60*(2), 121-150. doi: 10.1016/S0301-0511(02)00045-5

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449. doi: 10.1111/1467-9280.00383

Peng, S. H. (2000). Lexical versus 'phonological'representations of Mandarin Sandhi tones *Papers in Laboratory Phonology V: Acquisition and the Lexicon (Vol 5)* (pp. 152-167).

Qian, Z.-Y. (Ed.). (1997). *'Jinan Fangyan Cidian' (Jinan Dialect Dictionary)*. Nanjing: Jiangsu Education Press.

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., & Münte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia, 40*(7), 778-787. doi: 10.1016/S0028-3932(01)00188-9

Sakamoto, Y., & Ishiguro, M. (1986). *Akaike Information Criterion Statistics.* doi:10.1080/00401706.1989.10488538

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and speech, 34*(2), 145. doi: 10.1177/002383099103400202

Singmann, H. (2014). afex: Analysis of Factorial Experiments. R package version 0.10-113. http://CRAN.R-project.org/package=afex.

Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America, 115*, 1791. doi: 10.1121/1.1687832

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final/t. *Journal of Memory and Language, 52*(3), 322-338. doi: 10.1016/j.jml.2004.11.004

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*(30), 12530. doi: 10.1073/pnas.0609927104

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 394. doi: 10.1037/0278-7393.25.2.394

Vance, T. J. (1987). *An introduction to Japanese phonology*: State University of New York Press Albany.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics, 32*(2), 251-276.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1-25. doi: 10.1016/S0749-596X(03)00105-0

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*, 597. doi: 10.1121/1.2141003

Wong, P., & Diehl, R. L. (2003). Perceptual normalization for inter-and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research, 46*(2), 413. doi: 10.1044/1092-4388(2003/034)

Wood, S. (2006). *Generalized additive models: an introduction with R*: CRC Press.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73*(1), 3-36. doi: 10.1111/j.1467-9868.2010.00749.x

Wu, J. (2011). *Formulating the identification of mandarin Tone2 and Tone3 in multi-dimensional spaces.* Paper presented at the International Congress of Phonetic Sciences (ICPhS 2011).

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience, 170*(1), 223-231. doi: 10.1016/j.neuroscience.2010.06.077

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Yuan, J. H., & Chen, Y. Y. (2014). 3rd tone sandhi in standard Chinese: A corpus approach. *Journal of Chinese Linguistics, 42*(1), 218-237.

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhang, Q., & Zhu, X. (2011). The temporal and spatial features of segmental and suprasegmental encoding during implicit picture naming: An event-related potential study. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2011.09.040

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*(1), 25-64. doi: 10.1016/0010-0277(89)90013-9

Appendix. Crucial stimuli

Prime			related target		unrelated target	
JM real word	SC Hr pseudo-word	SC Lr pseudo-word	JM word	relatedness	JM word	relatedness
声音 sound /ʂəŋ in (R-R)/[a]	绳银 /ʂəŋ in (Hr-Hr)/	绳引 /ʂəŋ in (Hr-Lr)/	图像 video /tʰu ɕiɑŋ (Hl-Lf)/	3.95	结果 result /tɕiɛ kuo (L-Hl)[b] /	1.8
飞机 airplane /fei tɕi (R-R)/	肥集 /fei tɕi (Hr-Hr)/	肥挤 /fei tɕi (Hr-Lr)/	大炮 canon /ta pʰau (R-Lf)/	3.8	各位 each person /kɤ uei (R-Lf)/	1.65
签收 sign for /tɕʰiɛn ʂou (R-R)/	前熟 /tɕʰiɛn ʂou (Hr-Hr)/	前手 /tɕʰiɛn ʂou (Hr-Lr)/	快递 express delivery /kʰuai tʰi (R-Lf)/	4.45	馒头 steamed bun /man tʰou (R-H)[b]/	1.3
先驱 pioneer /ɕiɛn tɕʰy (R-R)/	咸渠 /ɕiɛn tɕʰy (Hr-Hr)/	咸取 /ɕiɛn tɕʰy (Hr-Lr)/	革命 revolution /kɤ miŋ (Hl-Lf)/	4.4	帮忙 help /paŋ maŋ (L-Hf)[b]/	1.15
回家 return home /xuəi tɕia (Hl-R)/	灰荚 /xuəi tɕia (Hl-Hr)/	灰甲 /xuəi tɕia (Hl-Lr)/	过年 Chinese New Year /kuo niɛn (Lf-Hf)/	4.3	领先 leading /liŋ ɕiɛn (Hl-R)/	1.2
流失 outflow /liou ʂʅ (Hl-R)/	溜十 /liou ʂʅ (Hl-Hr)/	溜史 /liou ʂʅ (Hl-Lr)/	水土 land and water /ʂui tʰu (Hl-Hl)/	4.45	快递 express delivery /kʰuai tʰi (R-Lf)/	1.5
无非 nothing more than /u fei (Hl-R)/	乌肥 /u fei (Hl-Hr)/	乌匪 /u fei (Hl-Lr)/	就是 just like /tɕiu ʂʅ (R-Lf)/	3.75	休息 rest /ɕiu ɕi (Lf-L)[b]/	1.1
有些 some /iou ɕiɛ (Hl-R)/	优鞋 /iou ɕiɛ (Hl-Hr)/	优写 /iou ɕiɛ (Hl-Lr)/	许多 many /ɕy tuo (Hl-R)/	3.45	技术 technique /tɕi ʂu (Lf-L)[b] /	1.95
大家 everyone /ta tɕia (Lf-R)/	大荚 /ta tɕia (F-Hr)/	大甲 /ta tɕia (F-Lr)/	各位 each person /kɤ uei (R-Lf)/	4.45	图像 image /tʰu ɕiɑŋ (Hl-Lf)/	1.25
拜托 please /pai tʰuo (Lf-R)/	拜驼 /pai tʰuo (F-Hr)/	拜妥 /pai tʰuo (F-Hr)	帮忙 help /paŋ maŋ (L-Hf)[b]/	4.65	革命 revolution /kɤ miŋ (Hl-Lf)/	1.05

Prime			related target		unrelated target	
JM real word	SC Hr pseudo-word	SC Lr pseudo-word	JM word	relatedness	JM word	relatedness
信息 information /ɕin ɕi (Lf-R)/	信席 /ɕin ɕi (F-Hr)/	信喜 /ɕin ɕi (F-Lr)/	技术 technique /tɕi ʂu (Lf-L)[b]/	4.45	就是 just like /tɕiu ʂʅ (R-Lf)/	1.5
肉汁（儿）gravy /ʐou tʂər (Lf-R)/	肉直（儿） /ʐou tʂər (F-Hr)/	肉纸（儿） /ʐou tʂər (F-Lr)/	馒头 steamed bun /man tʰou (R-H)[b]/	3.8	首先 first /ʂou ɕiɛn (Hl-R)/	1.05
到家 arrive home /tau tɕia (Lf-R)/	到荚 /tau tɕia (F-Hr)/	到甲 /tau tɕia (F-Lr)/	休息 rest /ɕiu ɕi (Lf-L)[b]/	4.35	许多 many /ɕy tuo (Hl-R)/	1.15
原因 reason /yæn in (Hl-R)/	冤银 /yæn in (Hl-Hr)/	冤引 /yæn in (Hl-Lr)/	结果 result /tɕiɛ kuo (L-Hl)[b]/	4.7	大炮 canon /ta pʰau (R-Lf)/	1.2
第一 first /ti i (Lf-R)/	第移 /ti i (F-Hr)/	第以 /ti i (F-Hr)/	首先 first /ʂou ɕiɛn (Hl-R)/	4.6	水土 land and water /ʂui tʰu (Hl-Hl)/	1.05
上风 windward /ʂaŋ fəŋ (Lf-R)/	尚缝 /ʂaŋ fəŋ (F-Hr)/	尚讽 /ʂaŋ fəŋ (F-Hr)/	领先 leading /liŋ ɕiɛn (Hl-R)/	3.7	过年 Chinese New Year /kuo niɛn (Lf-Hf)/	1

a. Abbreviation for tones. SC: Hl = High-level (1), Hr = High-rising (2), Lr = Low-rising (dip tone) (3), F = Falling (4). JM: R = Rising (1), Hf = High-falling (2), Hl = High-level (3), Lf = Low-falling (4), H = High (only in sandhi), L = Low (only in sandhi)

b. Tone Sandhi in the table: R+Hf -> L-Hf, R+Hl -> L-Hl, Lf+LF -> Lf-L/R-Lf, Hf+neutral -> R-H, R+neutral -> Lf-L

3 Are Tonal Realizations in One Dialect Predictable from Tonal Categories in the Other?

Abstract

Pronunciation dictionaries are usually expensive and time-consuming to prepare for the computational modeling of human languages, especially when the target language is under-resourced. Northern Chinese dialects are often under-resourced but used by a significant number of speakers. They share the basic sound inventories with Standard Chinese (SC). Also, their words usually share the segmental realizations and logographic written forms with the SC translation equivalents. Hence the pronunciation dictionaries of northern Chinese dialects could be easily available, if we were able to predict the tonal realizations of the dialect words from the tonal information of their SC counterparts. This paper applies statistical modeling to investigate the tonal aspect of the related words between a northern dialect, i.e. Jinan Mandarin (JM), and Standard Chinese (SC). Multi-linear regression models were built with between-word pitch distance of JM words as the dependent variable and the following were included as the predictors: SC tonal relations, between-dialect tonal identity, and individual backgrounds. The results showed that tonal relations in SC and between-dialect identity, as predictors featuring the relation between the JM and SC tonal systems, are significant and robust predictors of JM tonal realizations. The speakers' sociolinguistic and cognitive backgrounds, together with the tonal merge and neutral tone information within JM, are important for the prediction of JM tonal realizations and affect the way that between-language predictors take effect.

3.1 Introduction

3.1.1 The necessity and sufficiency of modeling under-resourced northern Chinese dialects

Under-resourced languages, featured by the 'lack of a unique writing system or stable orthography, limited presence on the web, lack of linguistic expertise, and lack of electronic resources for speech and language processing' (Besacier, Barnard, Karpov, & Schultz, 2014 p. 27), have always been a challenge for both engineers of Human Language Technologies (HLT) and linguists. One of the main reasons behind this challenge is the large amount of phonetic data required, which can be both difficult and expensive to acquire. To tackle this challenge, more and more researchers are transferring information from a related language or dialect to improve the understanding and automatic machine-processing of the under-resourced language. For instance, the automatic speech recognition of Afrikaans was significantly improved using the available Dutch data (Imseng, Motlicek, Bourlard, & Garner, 2014). However, to better incorporate the information from the related language, we need a better understanding about the relations between the two languages or dialects. In this aspect, linguists have carried out studies of a widerange of languages, though linguistic knowledge sometimes needs adaptations to be applied in engineering.

Chinese appears to be anything but under-resourced. For instance, Mandarin Chinese and Shanghai Chinese are already covered by the standardized multilingual text and speech database 'GlobalPhone' (Schultz, 2002). Even the (Standard) Mandarin-English bilingual test-to-speech system has seen important breakthroughs (Y. Qian & Soong, 2012). However, compared with the relatively well-investigated Standard Chinese (also referred to as 'Mandarin Chinese', 'Standard Mandarin', or 'putonghua', abbreviated as 'SC' in this article), many Chinese dialects are still under-resourced, including most northern dialects[2]. These northern dialects need more attention. First, they are used by a large Chinese population in everyday life (Hamed, 2005; R. Li, 1988). Second, they are closely related to SC and are usually used together with SC. This type of bilingualism comes with frequent code-switching/-mixing and sometimes also results in accented SC speech, which presents challenges for engineers and linguists (Huang, Chang, Zhou, & Lee, 2000; Sproat et al., 2004).

On the other hand, the close relation between the northern dialects and SC is also an attractive resource for the modeling of these dialects. Besides the large overlap in syntactic structure, the northern dialects and SC are very similar in basic sound inventories. For instance, we can find the comparison of the basic sound inventories of major Chinese dialects in a dictionary designed by linguists (Collective_work, 1989). This type of similarity has been proved useful in the sound-to-phoneme modeling in other languages (Imseng et al., 2014; Kamper et al., 2012; Van Heerden et al., 2010). However, there is one additional aspect of the between-dialect relation that may be useful and needs some more exploration. The northern dialects and SC share a high percentage of cognates and frequently borrow from each other[3] (Norman, 2003). The resulting translation equivalents share the same meaning across dialects and sound similar to each other. These related words are easy to identify because they are written in the same characters across all these dialects using the same logographic writing system. This paper applies statistical modeling to explore the tonal aspects of the related words between a northern dialect and SC. As a preliminary but important step before predicting the dialect pronunciation directly from SC pronunciation, the current study investigated to what extent and in what way a very limited but well available SC resource, the SC tonal categories, can predict the dialectal tonal realizations. We also tried to find out how the SC tonal categories, together with the speaker's social and cognitive backgrounds can account for the speaker-dependent tonal variability.

---

[2] The term 'northern dialects' is sometimes distinguished from 'Mandarin dialects', which are even more similar to Standard Chinese (Hamed, 2005). Here we use it in a more general way, following Li (1988).

[3] However, the cognates and loan words are difficult to distinguish for closely related dialects.

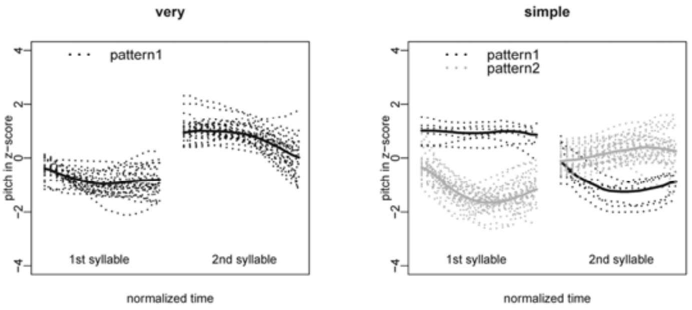
3.1.2 Research background on JM

We aim at predicting between-word pitch distances for JM Chinese using the tonal relations of the SC counterparts of the target words. JM is a northern dialect of Chinese. It is used in some local TV shows, but mostly in traditional folk arts, such as in 'Shandong Kuaishu'. Most JM speakers also speak SC fluently, and the mutual intelligibility between JM and SC is high (Tang & van Heuven, 2009). Some linguistic descriptions are available for JM. 'Jinan Fangyan Cidian' (JM Dialect Dictionary) (Z.-Y. Qian, 1997) provides the largest vocabulary but no recording. 'Jinanhua Yindang'(The Sound System of JM Dialect) (Z.-Y. Qian & Zhu, 1998) provides recordings of 428 monosyllabic characters, 410 words with two or more syllables, and some sentences. Pronunciations of characters are also available in 'Hanyu Fangyin Zihui' (Collective_work, 1989). However, these studies are based on the pronunciations by senior speakers many years ago (above 65 years old in 1993, 1998, and 1979).

Our fieldwork in 2012 showed that JM has become more similar to SC and the differences are mainly only retained in the tonal system. First, the usage and knowledge of JM-specific words are largely reduced and JM-specific words are replaced with words with etymologically related SC counterparts. Second, most JM words are now almost identical to their SC counterparts in segmental structure. However, the tonal differences remain between the JM and SC translation equivalents.

As a result, the current JM dialect enjoys a high percentage of related words with SC, which are almost only different from their SC counterparts in their tonal realizations (pitch contours). Since most non-tonal resources can already be directly transferred from SC, tone is the main potential space for cost reduction when building the pronunciation dictionary. The building cost of a JM pronunciation dictionary could be reduced if we are able to predict the tonal realizations of the JM words from the tonal information of their SC counterparts.

However, many JM words have shown tolerance of different tonal patterns, possibly due to the on-going process of 'lexical diffusion', where new tonal variants have appeared on some words but not on other words originally from the same tonal category (Chen and Wang, 1975; Wang, 1969), and the generalization of JM 'neutral tone sandhi' (Z.-Y. Qian, 1997), which means some words which were not reported to carry neutral tones are starting to have variants with neutral tone sandhi. As a result, some JM words allow one single tonal pattern (mono-pattern) but the others allow more than one (dual-pattern/ multi-pattern). Figure 1(a) and (b) demonstrate the difference between mono-pattern (i.e. 'very', fei1chang2, /feitʂʰaŋ/) and dual-pattern (i.e. 'simple', jiandan, /tɕiɛntan/). These words were plotted with normalized F0 contours from multiple speakers. Different tonal patterns of the same word can be observed not only in the production of different speakers but also in the production of the same speaker. This adds difficulty to the modeling of JM tonal realization.



(a) Acoustic variability (single-pattern, non-contrastive) (b) Tonal pattern variability (double-pattern, non-contrastive)

Figure 1. *Examples of mono-pattern (a) and dual-pattern (b) JM words, adapted from (Wu et al., 2014).*

3.1.3 Systematic correspondence and phonological similarity

In the present study, we try to implement two mechanisms which linguists have long been aware of in the modeling. The first mechanism is called ‘systematic correspondence’ (Dyen, 1963; Meillet & Ford, 1967). For two related dialects, ‘there is a significant number of words with similar meanings whose phonemes correspond systematically’ (Dyen, 1963 p. 634). Systematic correspondence can be measured via cross-dialectal comparison (M. Chen, 1973). Since the phonology of SC is based on the Beijing Mandarin (BM) pronunciation, we calculated the correspondence of the 4 JM tones and the 4 BM tones from the phonological transcriptions of 2,722 monosyllabic Chinese characters collected in 1979 (Collective_work, 1989), with polyphones counted multiple times[4], resulting in 3,679 pairs. Table1 shows the percentage of pairs, which follow the systematic correspondence rule. For instance, within the items carrying Tone 1 (high-level) in BM, 81% carry the low-rising tone in JM. The strength of systematic correspondence is likely to be even higher in current JM[5].

---

[4] For instance, according to the corpus the character for the Chinese quantifier ‘ge’ allows both Tone1 and Tone2 in Beijing and both low-rising and low-falling tone in Jinan. The resulting pairs would be the Cartesian product of the BM and JM sets, with an amount of four.

[5] However, how to handle polyphones and individual variability in choosing tonal patterns needs to be considered more carefully in implementation.

Table 1 Percentage of characters, which follow systematic correspondence in JM and Beijing Mandarin (BM, which the phonological system of SC is based on)

BM	Tone 1 high-level	Tone 2 high-rising	Tone 3 low-rising or dipping	Tone 4 high-falling	Total
JM	81% (low-)rising	76% high-falling	70% high-level	75% low-falling	76%

The second mechanism is ‘phonological similarity’. Phonological similarity across dialects or languages comes from two resources. (1) Cognates are inherently similar to each other. For instance, the Dutch ‘donder’ sounds like its English cognate ‘thunder’ because they were derived from the same proto-Germanic word *thunraz (Harper, 2001). (2) Loan words and their resources are similar to each other too. For instance, the Dutch ‘computer’ sounds identical to its English translation equivalent ‘computer’ because Dutch borrowed this word from English. Cognates and loanwords are difficult to differentiate when two dialects are involved, because most of the time the borrowed form is also a cognate. Compared with the differentiation based on word origin, the degree of between-language similarity is a more practical standard. In the relation between JM and SC, since the JM-SC related words are almost always with the same segmental structure, the most significant differences are in tone. If we keep minor pitch variation on acoustic level out of consideration, the related words are either identical or different in their tonal realizations. For instance, the disyllabic Chinese word ‘thanks’ (xie4xie5) /ɕiɛɕiɛ/ carries the falling+(low) neutral tones in SC and can have an almost identical tonal realization in JM, while ‘very’ (fei1chang2) /feitʂʰaŋ/ carries high-level+rising tones in SC but has a totally different low+high falling tone in JM. In the present study we distinguish only whether the JM word sounds identical to its SC counterpart. The between-dialect identity is taken as another main predictor in the current study.

Although both mechanisms are potentially useful for predicting the tonal realization of a JM word from its SC counterpart, the tonal phonological similarity can disrupt the effect of tonal systematic correspondence. For instance, as shown in Figure 1 (b), the disyllabic word ‘need’ in JM can follow the systematic correspondence rule with SC; then from the known tonal category of its SC counterpart (Tone 1 + Tone 4) we can predict that the JM ‘need’ (xu1yao4) /ɕyiau/ carries low-rising+low-falling tones, different from the high-level+falling tones in SC. On the other hand, the JM word ‘need’, can also carry tones almost identical to the tonal realization of its SC counterpart (with high-level+falling tone); in this case it disrupts the systematic correspondence rule between JM and SC. How the two mechanisms interact with each other in detail remains an open question.

Although the effects of systematic correspondence and phonological similarity have received full attention from linguists since the time of ‘Grimm’s Law’ (early 19th century), more efforts are still needed to bridge the gap between the linguistic theories and the technical applications. In the present research, we try to incorporate both mechanisms in the same statistical model and investigate their potential applications in predicting JM tonal realization using SC data.

3.1.4 Disyllabic tonal combinations and sandhi

We target disyllabic words in the present study. The majority of modern Chinese words are disyllabic, taking up 56% in the Microsoft Chinese dictionary (A. Wu & Jiang, 2000). Although in some Chinese dialects trisyllabic tonal realizations cannot be directly predicted from the corresponding disyllabic tonal sandhi patterns, e.g. in Tianjin Mandarin (Q. Li & Chen, 2012), disyllabic foot is the most frequent, the least constrained, and the standard foot in Chinese speech prosody (Feng, 2001; A. Li, 2002). This means the results from disyllabic words can be applied on multisyllabic words because they are usually realized with combinations of disyllabic and monosyllabic feet.

In SC, tonal realizations of disyllabic words are largely predictable from the citation tones of their monosyllabic components, either via tonal co-articulation or via sandhi rules (Y. Xu, 1994). Tone sandhi means the morpheme in combination carries a tonal variant different from the variant it carries in isolation. The Tone 3 sandhi in SC is well-known. Tone 3 before another Tone 3 sounds like Tone 2 but retains its minor acoustic difference from Tone 2 (Peng, 2000; Yuan & Chen, 2014). Tone 2 and the allophonic variants of Tone 3 have also been shown to be processed differently during speech preparation (Yiya Chen, Shen, & Schiller, 2011). In speech production and visual lexical access both the sandhi form of Tone 3 can activate and be activated by both Tone 2, which overlaps with its pitch contour, and Tone 3, which overlaps with its phonemic representation (Nixon, Chen, & Schiller, 2014). We have marked tonal categories of SC words according to the tonal citation forms and included whether the SC tonal categories of a pair of words are the same as one of the predictors.

It is important to note that the same combination of monosyllabic morphemes in JM can yield different disyllabic tonal realizations and the resulting variation is not totally predictable from the known predictors. According to Qian's (1997) description, JM has two types of sandhi. One type, the so-called 'normal' tonal sandhi maintains the distinctions across monosyllabic citation tones, and the tonal realizations of the disyllabic word need to be predicted from the tones of the citation forms of both syllables (Z.-Y. Qian, 1997). The second type concerns neutral tone. The neutral tonal sandhi merges the tonal realizations of different tones on the second syllable of a disyllabic word, so that the pitch contour of the disyllabic word can be predicted from (but not necessarily identical to) the citation tone of the first syllable. For instance, 'hen' (mu3ji1) /muʨi/ (with high-level+low-rising tone) and 'morning' (zao3shang4) /ʦauʂaŋ/ (with high-level+high-falling tone) are different in the second citation tone, but both can be realized with a 'low + high-level' tonal contour following the neutral tone sandhi rule. The JM neutral tone has different variants depending on the different citation tones of the previous syllables. For instance, the JM neutral tone can be realized as a low-falling tone following the low-rising citation tone, as a high-level tone following the high-falling or the high-level citation tone, and as a high-falling tone following the low-falling citation tone. Moreover, different from SC, the tonal realization of the syllable before the neutral tone is also different from its citation form. For instance, before the neutral tone, the rising citation tone is realized as a falling tone, the high-falling citation tone is realized as a rising tone, the high-level citation tone is realized as a low tone, and the low-falling citation tone is realized as a high-level tone (Z.-Y. Qian, 1997).

Most JM words carrying neutral tone sandhi have counterparts in SC, which also carry neutral tone. However, in our corpus, we have observed some JM words which allow tonal patterns following both neutral tonal sandhi and ‘normal’ tonal sandhi rules. For instance, the two patterns of ‘simple’ (jian3 dan4) /tɕiɛntan/ in Figure 1 (b) follow the two types of sandhi rules, respectively. Considering this phenomenon, whether the JM word in the specific rendition carries a neutral tonal sandhi or ‘normal’ sandhi is included as a predictor in the present study.

Note that the JM neutral tone sandhi with ‘high-level’ citation tone on the first syllable usually results in sandhi forms almost identical to its SC counterpart. For instance, the JM ‘snack’ (dian3xin5) /tiɛnɕin/ (with high-level+neutral -> low+high-level) sounds almost identical to the ‘snack’ (dian3xin1) /tiɛnɕin/ (with dip+high-level -> low+high-level) in SC. Thus some JM words identical to their SC counterparts result from JM neutral tone sandhi, not necessarily from borrowing.

3.1.5 Potential merging of JM tonal categories

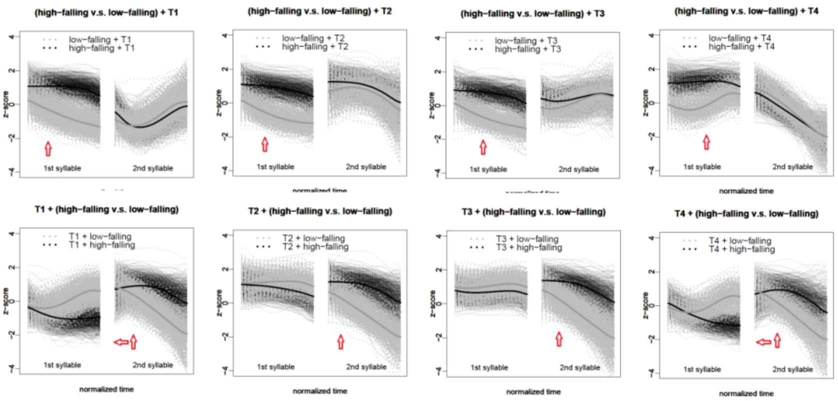
According to earlier descriptions, the high-falling and low-falling tones are similar in JM. We analyzed the monosyllabic words recorded in 1998 (Z.-Y. Qian & Zhu, 1998) and found the distributions of pitch contours of the two tones were very similar but still distinguishable in monosyllabic words. In our more recent corpus, the two tones are clearly distinguishable in disyllabic words, as shown in Figure 2.

Moreover, we found that some middle-aged speakers weaken the falling part of JM high-falling tone. As a result, the realization of JM high-falling tone is instead more similar to the JM high-level tone and maintains its contrast with the JM low-falling tone. This change holds for both monosyllabic and disyllabic words. As shown in Figure 3, the high-falling and high-level tones are very similar and the merge is more salient in the non-final position.

As shown in Figure 4, the difference between the JM low-rising and low-falling tones is also largely reduced when they appear in the non-final position, except when the following tone is a low-rising one.

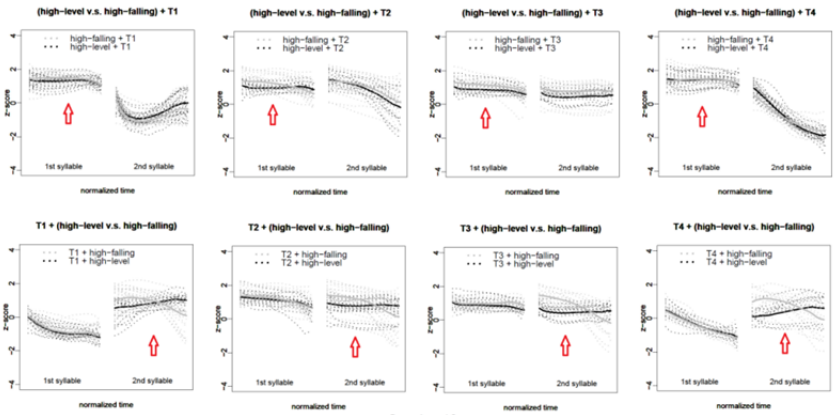
Due to the sandhi rules, the disyllabic combination of JM high-level+neutral tone realizes as low+high, very similar to the tonal realizations of JM low-rising+high-level, low-falling+high-level, low-rising+high-falling and low-falling+high-falling. The disyllabic combination of JM low-falling+neutral tone realizes as high-level+falling, very similar to the tonal realization of JM high-level+low-falling. These are depicted in Figure 5.

The above-mentioned types of potential merging are marked and taken into consideration in our modeling.



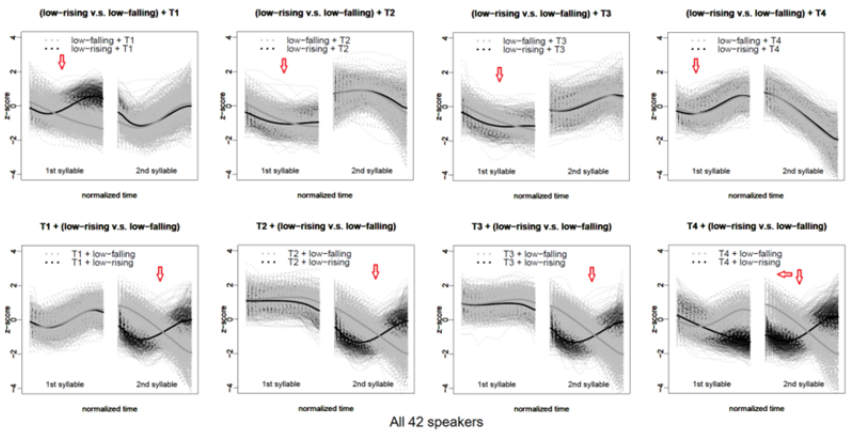
All 42 speakers

Figure 2. *The comparison of the JM high-falling and low-falling tones in the first syllable (first row) and the second syllable (second row) in disyllabic words.*



Speaker 18

Figure 3. *This speaker (Speaker 18) merges JM high-falling and high-level tones in the first syllable (first row) and almost merges them in the second syllable (second row) in disyllabic words.*



(low-rising v.s. low-falling) + T1

(low-rising v.s. low-falling) + T2

(low-rising v.s. low-falling) + T3

(low-rising v.s. low-falling) + T4

T1 + (low-rising v.s. low-falling)

T2 + (low-rising v.s. low-falling)

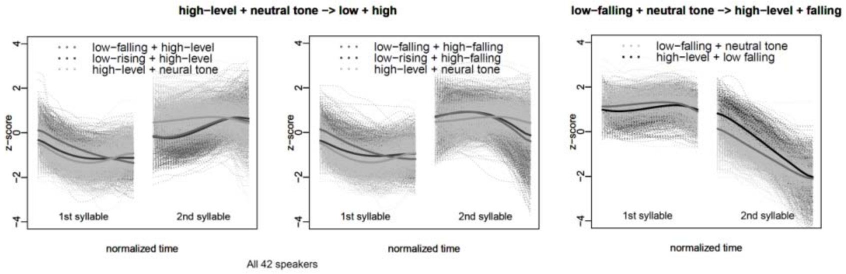
T3 + (low-rising v.s. low-falling)

T4 + (low-rising v.s. low-falling)

All 42 speakers

Figure 4. *The difference between JM low-rising and low-falling tones is largely reduced in the first syllable (first row) but the difference is maintained in the second syllable.*

.



high-level + neutral tone -> low + high

low-falling + neutral tone -> high-level + falling

All 42 speakers

Figure 5. *JM high-level+neutral tone realizes as low+high, very similar to the tonal realizations of JM low-rising/low-falling+high-level and low-rising/low-falling+high-falling.*

3.1.6 Word frequency

The frequency effect has long been known and discussed by psycholinguists in research on lexical access (Dell, 1990; Grainger, 1990; Levelt et al., 1999; Oldfield and Wingfield, 1965). The general finding is that frequent words and forms are accessed more quickly.

We would like to know whether word frequency affects the tonal relation across JM words and whether word frequency modulates the tonal effects of the systematic correspondence between SC and JM. Moreover, we would like to see whether including word frequency information will improve the modeling of JM tonal realizations.

Limited by the sharing of logographic writing system in Chinese, we have no access to any dialect-specific word frequency data. We use the Chinese word frequency based on film subtitles (Cai & Brysbaert, 2010). This resource can be taken as mainly in SC, but we cannot exclude the contribution of JM speakers.

3.1.7 Individual backgrounds

Individual variation is interesting for both linguists and technology experts. System developers working on speaker adaptation have been pretty successful in dealing with pure acoustic deviations via speech normalization and changing Hidden-Markov-Model (HMM) parameters (Leggetter & Woodland, 1995; Woodland, 2001). However, when separate models are built for different speaker types and/or when the pronunciation dictionary also needs to be adapted for different accents, the cost increases and the speaker type is difficult to decide (Huang et al., 2000; Woodland, 2001). On the other hand, sociolinguists have proven that these phonological variations can largely be predicted from socio-backgrounds (Labov, 2006; Weinreich, Labov, & Herzog, 1968). Moreover, socio-backgrounds can not only group the speakers but can also index the speakers along a continuum. It would be beneficial if measurable individual backgrounds are introduced into the model.

Earlier studies reported segmental variation across and within JM individuals, regarding age and speech style (Cao, 1991; Qian, 1997; Qian and Zhu, 1998). They distinguished the ‘old’ JM from the ‘new’ JM and the ‘Wendu’ (literal style) from the ‘Baidu’ (colloquial style). They also did some quantitative analyses. However, these studies did not pay attention to the relation between tonal variation and individual backgrounds and after 20 years the JM words’ segmental structures are mostly identical to that of their SC counterparts.

We intend to incorporate individual backgrounds into the predicting model and investigate their statistical effects on JM tonal realization. Our corpus was collected in 2012. It covers a greater age range of urban native JM speakers and includes individual backgrounds on both social and cognitive aspects. We expect to quantify the age effect observed earlier (Z.-Y. Qian, 1997) and we are also interested in which of the other aspects take effects.

Beside the previously investigated factors, such as gender, age, and education backgrounds, we also take the speakers’ experience with both languages into consideration. Previous studies on bilingualism have shown that language of education and language proficiency affect the pattern of code-switching in bilingual speech production (Carter, Davies, Parafita Couto, & Deuchar, 2010). Whether these factors influence the systematic correspondence in general needs further investigation. Language proficiency is influenced by language exposure. Hence frequencies of language usage were taken into consideration. Beside sociolinguistic backgrounds, the cognitive aspects may also affect individual variation of speech production. For instance, tonal awareness, as a subset of phonological awareness, reveals listeners’ aptitudes for discriminating and identifying tones (X. Chen et al., 2004; Shu, Peng, & McBride-Chang, 2008). It also affects the processing of tonal variants in JM lexical access (J. Wu & Chen, 2014). Additionally, the speaker’s digit-naming speed and auditory working memory are also taken into consideration. By far, few studies have shown relevance of auditory working memory in speech production. Even bilinguals seem to be similar to monolinguals in auditory working memory (Bialystok, Craik, & Luk, 2008) and simultaneous interpreters seem to have no advantage in retaining auditory information (Signorelli, Haarmann, & Obler, 2011). However, would auditory working memory influence the systematic correspondence shown in the bilinguals' production? All these factors are taken into consideration, together with their interaction with age.

It is reasonable to assume that the effects of cognitive and socio-linguistic backgrounds are at least partly mediated by age in the present study. Age is easy to measure. However, age is related to both the aging of the individuals and the change of the society. On the one hand, cognitive aging affects the speakers' cognitive performances. Older speakers have declined auditory working memories and slower reaction times. On the other hand, socio-linguistic backgrounds change across generations and affect the change of the tonal system. With the promotion of SC and social progress in China, younger speakers use more SC and less JM, receive higher education, and are more likely to receive their literacy educations in SC. The change of socio-linguistic backgrounds can also influence some cognitive backgrounds. With the introduction of the alphabetic system 'pinyin', the younger generations received more training in benefit of their acuity to tones.

Nevertheless, it is also reasonable to hypothesize that the cognitive and socio-linguistic backgrounds could affect JM tonal realization beyond the effect of age. Individuals of the same age have different cognitive aptitudes and individuals from the same generation have different socio-linguistic experiences. Which cognitive and socio-linguistic backgrounds have unmediated effects on the JM tonal system?

3.1.8 Research predictions

The present study focuses on the influences of the following factors on JM tonal realization: systematic correspondence, phonological similarity, and individual backgrounds. Other covariates are also taken into consideration, including the neutral tone sandhi and potential tonal merging in JM and word frequency.

As mentioned above, the systematic correspondence is related to the between-word tonal relations in both dialects. The between-word tonal relation, whether measured in a scale or binomially, is comparable across different tonal categories. Thus, answers about the between-word tonal relation can be technically applied before identifying the specific tonal categories. Also, the answers to the theoretical questions are category-independent and more general. We choose between-word pitch distance as the dependent variable to quantify the effects and interactions of the above-mentioned factors. Based on the linguistic knowledge of systematic correspondence (described in 1.3), we expect them to affect the between-word pitch distance in JM in the following way.

The systematic correspondence mechanism predicts that, if two words share their tonal categories in one dialect, their counterparts are also more likely to share tonal categories in the other dialect. Sharing tonal categories means smaller pitch distance. Thus, considering the pitch distance between two JM words, the distance is more likely to be smaller if their counterparts share tonal categories in SC. JM disyllabic words whose SC counterparts share the tonal categories on both syllables should show smaller pitch distance compared to JM disyllabic words whose counterparts only share the tonal category on just one syllable.

The effect of systematic correspondence should be robust if neither of the two JM words is realized identically to its SC counterpart. However, the effect of systematic correspondence should be disrupted when one of the JM words borrows its tonal realization from SC, especially when the two words share tonal categories in SC. For instance, ‘direct’ (zhi2jie1) /tʂʅtɕiɛ/ and ‘leave’ (li2kai1) /likʰai/ share the same ‘Tone 2+Tone 1’ (high-rising+high-level) tones in SC. Systematic correspondence predicts that they are very likely to share the same ‘high-level + low-rising’ tones in JM and see a relatively small between-word pitch distance. However, if the speaker borrows the SC form of ‘leave’ (li2kai1) /likʰai/ (with high-rising+high-level) into JM and keeps the JM native high-level+low-rising tone for ‘direct’ (zhi2jie1) /tʂʅtɕiɛ/, the between-word pitch distance should be larger than expected.

On the other hand, the predicting power of the two words’ tonal relation in SC should be rebuilt when both of the JM words borrow their tonal realizations from SC. Moreover, the predicting power should be stronger, because the effect is no longer mediated by systematic correspondence, which does not control all of the JM vocabulary (see Table 1). In the present case, the pitch distance between two JM words directly reflects the pitch distance of their SC counterparts. For instance, if both JM ‘leave’ (li2kai1) /likʰai/ and ‘direct’ (zhi2jie1) /tʂʅ tɕiɛ/ borrow the SC form, it is sure that they share the high-level+low-rising tones just like in SC and have a small between-word pitch distance.

Note that, when a JM word is realized identically to its SC counterpart, it is difficult to decide from the surface realization whether it is due to borrowing or the coincidence of JM neutral tone sandhi (see 1.5). Nevertheless, whether the between-dialect identity is due to borrowing or the specific JM neutral tone sandhi, it should work similarly in most cases.

Also, the JM neutral tone sandhi itself should work in a similar way as the between-dialect identity. For instance, ‘body’ (shen1ti3) /ʂəntʰi/ and ‘clear’ (qing1chu3) /tɕʰiŋtʂʰu/ share the same high-level+low-rising (‘Tone1 + Tone3’) tones in SC. Systematic correspondence predicts that they are very likely to share the same low+high-level tones in JM and see a relative small pitch distance. However, ‘clear’ is usually produced with neutral tone sandhi in JM and realizes with a low-falling+low pitch contour. As a result, the between-word pitch distance should be larger than expected.

We are also interested in the effects of individual backgrounds. The present study does not focus on the pure physical gender differences which can be normalized, such as the gender effect on the pitch range (Chen, 2011; Peng et al., 2012). The present study works on the variants which simple normalization cannot handle, namely the JM tonal variants related to the speakers’ social backgrounds.

Except gender, most of the other aspects of individual backgrounds are related to age. Older speakers are more proficient in JM, use JM more frequently, and mostly received literacy education in JM. It is natural to predict that their JM pronunciation should be less aligned with SC. However, we know JM is related to SC via both systematic correspondence and phonological similarity. Does it mean that older JM speakers’ between-word pitch distance should be less sensitive to the words’ tonal relation in SC? Or does it mean that older JM speakers produce JM words acoustically less similar to SC? Or are both true? On the other hand, older speakers usually have lower auditory working memories due to aging, and lower tonal awareness because they didn't receive proper education of 'pinyin' (the Chinese alphabetic writing system). Do these cognitive predictors have independent effects besides age? We will apply statistical analyses to answer these questions.

3.2 Material and methods

3.2.1 Corpus preparation

The speech data used in the present study were collected from 42 JM native speakers in 2012 (see section 2.2 for details). Each speaker read 400 disyllabic Chinese words in JM. The written words were selected from a corpus of Chinese film subtitles (Cai & Brysbaert, 2010). One list of 200 high-frequency words was selected from the 10% disyllabic Chinese words with the highest word frequency. In a similar way, we selected the other list of 200 low-frequency words. In each list, there are 10 words for each of the 20 disyllabic tonal combinations. The high and low frequency lists were presented to the speakers in two blocks with a self-paced rest break in between. The words in each list were presented in a different random order for each speaker. After the speaker finished producing a word, they pressed a key to see the next word.

We used Praat (Boersma & Weenink, 2001) to extract pitch contours. Only pitch contours on the rhymes were extracted. A trained phonetician listened to each recording, looked at the spectrogram, and manually marked the rhyme of each syllable. Also, in this process, recordings with speech and recording errors were excluded from the corpus. Afterwards, the pitch contours were converted from hertz to semitones with 100Hz as the base and then transformed into z-scores based on the speakers' means and standard deviations (Chen, 2011; Lobanov, 1971). This normalization removed the pitch range difference across speakers, which is not the main focus of the present study. The normalized pitch contours were then interpolated to 20 points per-syllable to remove the difference in duration. Since each speaker produced a list with many different tones, the multidimensional distribution of the dataset involves inherent clusters. Thus, we chose a density-based local approach to eliminate the possible outliers (Breunig, Kriegel, Ng, & Sander, 2000). We calculated Local Outlier Factors (LOF) for each speaker's pitch contours. Any pitch contour with an LOF greater than 1.5 (Breunig et al., 2000) and belonging to the 2.5% with the highest integral density was eliminated from the corpus.

3.2.2 Individual backgrounds

We collected both sociolinguistic and cognitive backgrounds from the speakers. The sociolinguistic backgrounds included the speaker's age, gender, education level, reported proficiencies and frequencies of JM and SC, language of literacy education, and the dialects they use with their primary social relations. All speakers except one received formal education, of which 57% reached college level and the rest reached middle school level. As for the literacy education, 26% speakers received it in JM, 56% received it in SC, and 18% received it in a combination of JM and SC. The cognitive backgrounds were found by earlier studies to be related to vocabulary, reading, and comprehensive skills, including the speaker's digital naming speed (Torgesen and Davis, 1996) in JM and SC, auditory working memory in JM (Gathercole, Willis, Baddeley, & Emslie, 1994), and tonal awareness of JM and SC (Shu, Peng, & McBride - Chang, 2008). The distributions of the scale variables were plotted in Figure 6.

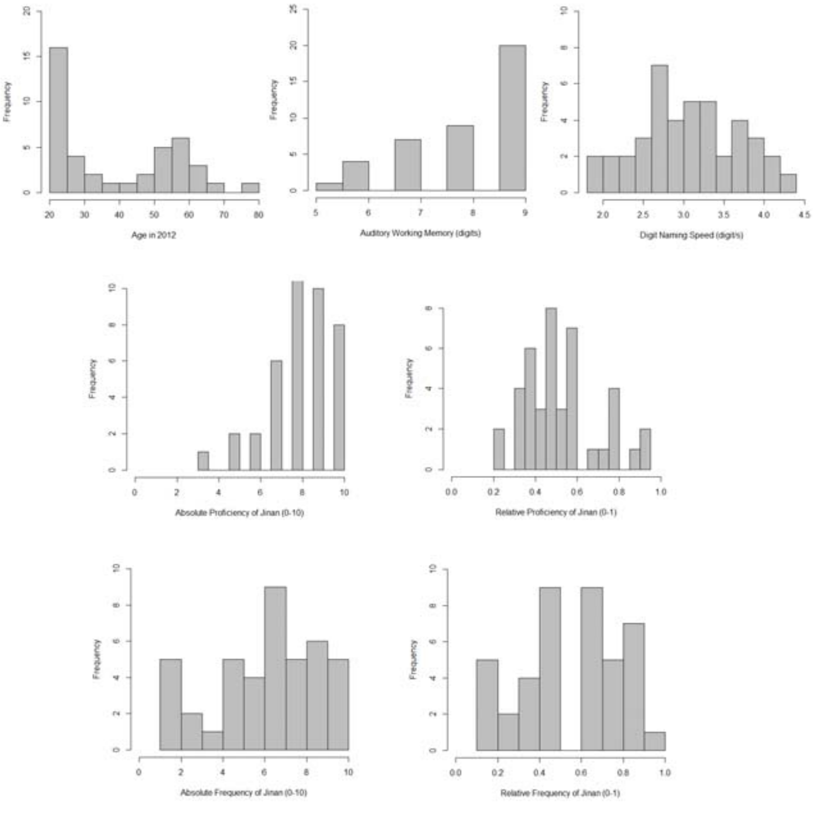


Figure 6. *The distribution of individual backgrounds: age (first row left), auditory working memory (first row middle), digital naming speed (first row right), absolute proficiency of JM (middle row left), relative proficiency of JM (middle row right), absolute frequency of JM (last row left), and relative frequency of JM (last row right).*

3.2.3 Model fitting

A 'between-word pitch distance' was used as the crucial dependent variable in the modeling. This was practically defined as the Euclidean distance (Deza & Deza,

2009) of the pitch contours between each two JM words produced by the same speaker. For each speaker, each normalized pitch contour was taken as a Euclidean vector. For each pitch contour, each of the 20 time points was taken as one dimension of the vector. Then Euclidean distances were calculated for each two vectors of the same speaker, yields 53,628 to 79,800 between-word pitch distances for each speaker. Similar acoustic distance matrix has been used in studies investigating the correlation between speech perception and production (Iverson et al., 2003).

The following two sets of predictors were included in the modeling following the research predictions. The first set includes linguistic predictors, which are nested with pairs of words. Table 2 shows the structures and explanations of these predictors. The SC tonal relations (on the first and second syllables) were calculated from the standard phonological transcriptions (pinyins) of SC words. A phonetician with Putonghua Proficiency Test Certificates- Level 1B judged whether the JM word was produced (almost) identically to its SC counterpart, taking native monolinguals' and bilinguals' similarity rating of a subset of the corpus in another study into consideration (Wu et. al. in prep.). Whether the two JM words are undergoing tonal merging and on which syllable(s) they are merging were predicted from the tonal categories of their SC counterparts first and then manually verified. Whether the JM word carries a neutral tone was predicted from the tonal category of its SC counterpart first and then manually verified. Word frequency was imported from the recording list (Cai & Brysbaert, 2010) and converted into two categories. The second set includes predictors of individual backgrounds, which are nested with speakers and were collected together with the recordings. Table 3 shows the structures and explanations of these predictors.

Table 2. Linguistic predictors

Predictor	Structure	Explanation
Tonal relation on the first SC syllables	2 levels	whether or not the counterparts of the first syllables are from the same tonal category in SC; an indicator of systematic correspondence
Tonal relation on the second SC syllables	2 levels	whether or not the counterparts of the second syllables are from the same tonal category in SC; and indicator of systematic correspondence
Between-dialect identity	3 levels	whether neither, one, or both of the two words is/are identical to its/their counterpart(s) in SC
Merge	5 levels	whether or not the tones of the two words are undergoing merging on neither, the second, the first, both, or the combination of the two syllables in JM
Neutral tone	2 levels	whether or not this pair involves neutral tones
Word frequency	3 levels	whether or not the two words are from different word frequency groups, both from the high frequency group, or both from the low frequency group

Table 3. Predictors of individual backgrounds

Predictor	Structure	Explanation
Gender	2 levels	male or female
Age	interval	the speaker's age in 2012
JM absolute proficiency	interval	self-rated JM proficiency on a 1 to 10 scale
JM relative proficiency	interval	the proportion of self rated JM proficiency in the speaker's total language proficiency
JM absolute frequency	interval	self-rated JM frequency on a 1 to 10 interval
JM relative proficiency	interval	the proportion of self-rated JM frequency in the speaker's total language frequency
Education	2 levels	the highest education the speaker has received (middle school, college)
Language of literacy education	3 levels	whether the speaker received literacy education in JM, SC, or a mixing of both dialects
Tonal awareness in JM	interval	the correct rate in JM tonal oddity test
Tonal awareness in SC	interval	the correct rate in SC tonal oddity test
Digit-naming speed	interval	digit-naming speed in JM (words/per second)
Auditory working memory	interval	how many digits the speaker can recall in correct order immediately after the digits are presented in JM

We performed exploratory linear-mixed-effects (LME) analyses on the between-word pitch distance data, using R (R_Core_Team, 2013), lme4 (Bates, Maechler, Bolker, & Walker, 2013), and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2013). Considering the size of the dataset, building all the predictors into one model would exceed the limitation of calculating power. Moreover, a model with too many predictors would suffer from multi-collinearity across predictors and yield unintelligible results. Thus, we did not build a model including all the predictors. Instead we built smaller models using different subsets of the data and subsets of the predictors to investigate the importance and robustness of different predictors. Then we built the important predictors and their interactions together in a more general model.

We fitted two sets of LME models. The first set of LME analysis focused on the effects of the SC tonal relation and the other linguistic predictors. We built one separate model for each speaker. Each speaker-wise model included all the fixed effects of the six nominal linguistic predictors in Table 2 (tonal relation on the first SC syllables, tonal relation on the second SC syllable, between-dialect identity, merge, neutral tone, word frequency), and their two-way and three-way interactions, as well as the random intercept of the SC tonal combinations of the two words. These models were trimmed and the results of the final models are reported here.

The second set of LME analysis focused on exploring the effects of individual backgrounds (see Table 3) and their interactions with the SC tonal relations predictor (a combination of the tonal relations on the first SC syllables and second syllables). To avoid unnecessary rank-deficiency and accommodate for the limit of calculating power, this analysis was performed on the averaged between-word pitch distance, which was collapsed across pairs and aggregated by the combination of speaker, SC tonal relation, and between-dialect identity. A separate model was built for each level of between-dialect identity because the speaker-wise models showed that SC tonal relations function differently with different levels of between-dialect identity. We did not build all the predictors of individual backgrounds at once in one model because of two main reasons: first, the collapsed data could not support so many predictors; second, the multi-collinearity between these predictors would blur the interpretation of mediated effects and result in a model with unclear causality and unreliable direction of main effects. Instead, we first included these predictors separately in smaller models to investigate their independent effects and then, after statistically removing the collinearity, built a selected subset of the predictors together in a full model to investigate their interactions and non-mediated effects.

In the separate analyses of individual backgrounds, each model included SC tonal relations, one aspect of the individual backgrounds (interval predictors centralized by subtracting the mean), and their interaction as the fixed predictors, as well as Speaker as the random intercept. However, in the combined analysis of individual backgrounds, in order to investigate the non-mediated effects of the individual backgrounds, all the interval predictors of individual backgrounds (Age, JM absolute proficiency, JM absolute frequency, digit-naming speed, tonal awareness in JM, and tonal awareness in SC) were first centralized and standardized. Then Kappa (R. H. Baayen, 2011; Belsley, Kuh, & Welsch, 2005) and pair-wise Pearson correlations were calculated for these interval predictors to quantify the problem of collinearity and the multicollinearity was reduced via residualization (R. Harald Baayen, Feldman, & Schreuder, 2006; Jaeger, 2010). The nominal predictors of individual backgrounds (Gender, Education, and Language of literacy education) were also considered for their multicollinearity and separate models were fitted accordingly on subsets of data to see whether the effects persisted.

All these models were fitted in an exploratory way. We first built full models, including all the predictors of interest and their two-way and three-way interactions as the fixed predictors, as well as the random intercept of the SC tonal combinations of the two words. When there were unrealized combinations of the predictors, which revealed multi-collinearity and would cause rank deficiency in the modeling, the corresponding interaction terms were removed. A backward elimination was then performed to remove non-significant effects, using p-values calculated from F test based on Sattethwaite’s method (Kuznetsova et al., 2013). Finally, we carried out post-hoc tests and calculated the least squares means and confidence intervals for the nominal predictors (Kuznetsova et al., 2013). As for the interval predictors and their interactions with the nominal predictors, we calculated and plotted the slopes of estimated mean distances.

3.3 Results and discussion

The speaker-wise models showed that the words' SC tonal relations, the between-dialect identity, their two-way and three-way interaction, and the SC tonal combination all affected JM between-word pitch distance. The proportion of variance accounted for by the final models (R^2) ranges between 0.25 and 0.60 (mean = 0.43, median = 0.44), indicating that the predictive power of these models ranges between medium and large. The analysis of individual backgrounds showed that the effects of SC tonal relations were modulated by the speakers' sociolinguistic and cognitive backgrounds. In every final model, the fixed predictor of the SC tonal combinations of the two words was kept, indicating that SC tonal combination was robust in predicting the between-word pitch distance in JM. In the following texts, results are reported and interpreted based on F statistics and post-hoc estimates of the models (Kuznetsova et al., 2013). Estimates yielded by the model summaries are averaged across all the individual models and reported in the appendix.

3.3.1 Systematic correspondence works: effects of SC tonal relation on JM between-word pitch distance

The effects of SC tonal relations reveal the effect of the systematic correspondence mechanism, which predicts that two words that share tonal categories in SC would have a smaller between-word pitch distance in JM.

In the speaker-wise models, the main effect of the tonal relation on the first SC syllables was significant for most speakers, $F \in [7.45, 522.19]$, $p < 0.05$, except for Speaker08, $F = 2.96$, $p = 0.08$ and Speaker20, $F = 3.84$, $p = 0.05$, while the main effect of the tonal relation on the second SC syllable was significant in 24 of the models, $F \in [4.55, 77.92]$, $p < 0.05$ but insignificant in 18 of the models, $F \in [0.00, 3.43]$, $p > 0.05$, indicating that the tonal relation on the first SC syllables was more robust than that on the second syllable as a predictor. The two-way interaction of the tonal relations on the first and second SC syllables was significant in 7 of the models, $F \in [4.31, 9.75]$, $p < 0.05$ but insignificant in 35 of the models, $F \in [0.00, 3.60]$, $p > 0.05$, indicating that the tonal relation in SC was relatively independent on the first and the second syllables in predicting JM between-word pitch distance. In Figure 7, the estimated means and confidence intervals from the speaker-wise models were plotted in clusters according to the conditions. The combination of the tonal relations on the first and second SC syllables [whether neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) of the SC counterparts are from the same tonal category] was represented with color-coded clusters and labels on the horizontal axis and the different levels of between-dialect identity were plotted in separate planes. As shown in Figure 7 (left and right planes), when two words shared the tonal category of their first or second syllables in SC, the between-word pitch distance of their counterparts in JM was also reduced. Sharing more tonal categories in SC reduced the between-word pitch distance in JM. Also sharing the tonal category on the first syllable reduces the distance more than sharing it on the second syllable. Although the effects of SC tonal relations were largely removed when only one of the two JM words was identical to its SC counterpart, in the next session we will discuss the causes.

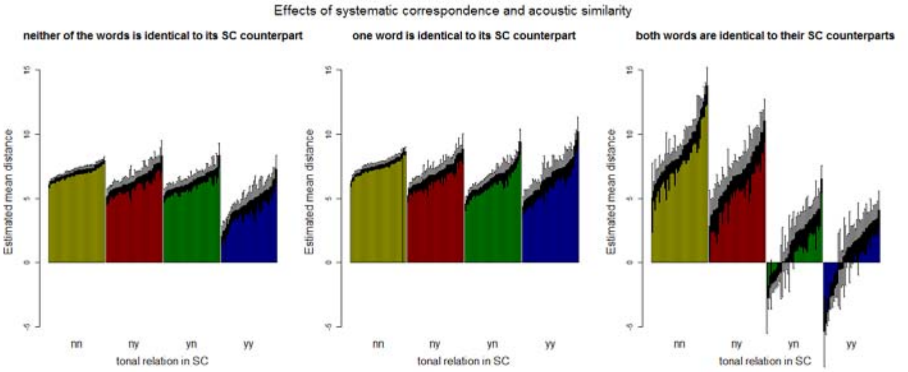


Figure 7. *The interaction of the tonal relation in SC and between dialect identity on the estimated between-word pitch distance of JM words. Individual estimates under the same condition were clustered according to tonal relation in SC [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category]. The estimates were split into three plots according to between dialect identity [neither (left), one (middle), or both (right) of the two words is/are identical to its/their counterpart(s) in SC].*

3.3.2 Phonological similarity interrupts and reinstalls systematic correspondence: effects of between-dialect identity

The phonological similarity mechanism predicts that, when one of the two JM words was realized as almost identically to its SC counterpart, the predicting power of systematic correspondence would be disrupted on this word. Under this condition, whether or not this word shares tonal categories with another word in SC would no longer predict the pitch distance between these two words, unless the other word is also realized as almost identically to its SC counterpart. The results fully support the theoretical predictions.

In the final speaker-wise models, the main effect of between-dialect identity was significant in all the models, F∈[5.36, 689.90], $p < 0.05$. However, the direction was inconsistent. More importantly, between-dialect identity interacts with SC tonal relations. The two-way interaction of the tonal relation on the first SC syllables and between-dialect identity was significant for all the speakers, F∈[7.10, 589.70], $p < 0.05$. The two-way interaction of the tonal relation on the second SC syllable and between-dialect identity was significant in 40 of the models, F∈[3.32, 127.10], $p < 0.05$ but insignificant in 2 of the models, F∈[1.34, 2.68], $p > 0.05$, indicating that this interaction was robust across speakers but less reliable than its counterpart involving the first syllable. The 3-way interaction of the tonal relation on the first SC syllables and second syllable and the between-dialect identity was significant in all of the models, F∈[3.19, 131.00], $p < 0.05$, except that the term was removed for

Speaker17 and Speaker26 due to missing combination. Looking back to Figure 7, when one of the JM words was identical to their SC counterparts (left plane), sharing tonal categories in SC reduced the between-word pitch distance in JM, revealing the power of systematic correspondence mechanism. A similar pattern was also found when both of the words were identical to their SC counterparts (right plane). In this situation the JM between-word pitch distance represents the between-word pitch distance of their SC counterparts[6]. However, when one JM word was identical to its SC counterparts and the other was not, the effects of SC tonal relations were largely removed. This pattern is consistent with the prediction from the phonological similarity mechanism.

Additionally, the difference in between-word pitch distance induced by different SC tonal relations was greater when both of the words were identical to their SC counterparts than when neither of the words was identical to its SC counterpart. This is reasonable. After all, the SC tonal relation is indirectly related to JM between-word pitch distance via JM tonal relation in the former case but is directly reflected by the SC between-word pitch distance in the latter case.

Similar to the results in the speaker-wise models, the main effect of the SC tonal relations was significant in all the models including individual backgrounds and it functioned differently with different levels of between-dialect identity. When neither of the two words was identical to its SC counterpart, sharing tones in SC reduced the JM between-word pitch distance and sharing tones on the second syllable reduced the distance more than sharing tones on the first syllable, $F\_neither \in [201.90, 1140.22]$, $p < 0.05$. Similarly, when both of the two words were identical to their counterpart in SC, sharing tones in SC also reduced the JM between-word pitch distance. However, under this condition, sharing tones on the first syllable instead of on the second syllable reduced more distance, $F\_both \in [23.11, 521.81]$, $p < 0.05$. When one of the two words was identical to its SC counterpart, this effect was reversed in that sharing tones in SC increased the JM between-word pitch distance, $F \in [21.86, 41,92]$, $p < 0.05$.

3.3.3 Neutral tone disrupts systematic correspondence

JM neutral tone generally disrupted the predicting power of systematic correspondence. In the speaker-wise models, the main effect of neutral tone was significant in 38 of the models, $F \in [6.64, 6106.00]$, $p < 0.05$ but insignificant in 4 of the models, $F \in [0.39, 1.85]$, $p > 0.05$. The direction of the main effect of neutral tone was inconsistent across speakers. The two-way interaction of the tonal relation on the first SC syllables and neutral tone was significant in 39 of the models, $F \in [7.53, 2002.00]$, $p < 0.05$ and insignificant in 3 of the models, $F \in [0.39, 2.96]$, $p > 0.05$. The two-way interaction of the tonal relation on the second SC syllable and

---

[6] Here the model estimated unrealistic negative values for some speakers when the SC counterparts are from the same tonal category in the first or both of the syllable(s), revealing some problem of over-fitting.

neutral tone was significant in 40 of the models, $F \in [3.93, 2829.00]$, $p < 0.05$, insignificant in 2 of the models, $F \in [0.38, 2.85]$, $p > 0.05$. The three-way interaction of the tonal relations on the first and second SC syllables and neutral tone was significant in 19 of the models, $F \in [3.87, 37.11]$, $p < 0.05$ and removed in 23 of the models. Although sharing the tone on the first or second SC syllable generally reduced the between-word pitch distance, JM neutral tones counterweighed these effects.

Taking between-dialect identity into consideration, we found more complex interactions. The two-way interaction of neutral tone and between-dialect identity was significant in 26 of the models, $F \in [4.35, 246.40]$, $p < 0.05$, insignificant in 4 of the models, $F \in [0.55, 2.78]$, $p > 0.05$, and removed in 12 of the models. The three-way interaction of the tonal relation on the first SC syllables, neutral, and between-dialect identity was significant in 19 of the models, $F \in [50.61, 307.71]$, $p < 0.05$ but removed in 23 of the models. The three-way interaction of the tonal relation on the second SC syllable, neutral, and between-dialect identity was significant in 10 of the models, $F \in [4.66, 54.44]$, $p < 0.05$ but removed in 32 of the models. The post-hoc analysis showed that neutral tones interacted with the SC tonal relations in different ways depending on the condition of between-dialect identity. When neither or only one JM word was identical to its SC counterpart, the involvement of neutral tone generally increased the between-word pitch distance and reduced the effect of SC tonal relations. This is consistent with the general finding that neutral tones disrupted the predicting power of systematic correspondence. However, when both of the two words were identical to their SC counterparts, neutral tones enhanced the effect of the tonal relations on the first SC syllables but reduced the effect of the tonal relations on the second SC syllables.

Why did neutral tones disrupt the effect of the SC tonal relation on the second syllable? This is probably because there is no unified realization of neutral tone and the pitch contour of the neutral tone depends on the tonal category of the preceding syllable. When two SC words both carry neutral tones on the second syllables, the two neutral tones can be realized as very different variants, so long as their previous syllables carry different tones. Thus sharing SC neutral tones cannot reduce the between-word pitch distance when both of the two JM words were identical to their SC counterparts. Similarly, when neither of the JM words was identical to their SC counterparts and the SC words carry neutral tones, the systematic correspondence mechanism predicts that the JM words are also more likely to carry neutral tones. The two JM neutral tones are not necessarily similar in pitch contours either. Hence unlike the case of sharing the other tonal categories, sharing neutral tones on the second syllable does not reduce the between-word pitch distance.

Why did neutral tones also disrupt the effect of the tonal relations on the first SC syllables when neither of the JM words was identical to its SC counterpart? When both second syllables carry neutral tones, the pitch contours and the between-word pitch distance depend on the tones of the first syllables. However, the same JM citation tone is realized as different sandhi forms before a neutral tone and before the other tones. As a result, when one of the second syllables carries a neutral tone and the other does not, sharing citation tones on the first SC syllables cannot reduce between-word pitch distance of the JM words.

Then why did neutral tones instead enhance the effect of the tonal relations on the first SC syllables when both words were identical to their SC counterparts? Unlike the JM tones, the SC tones preceding neutral tones realize very similar to the corresponding citation forms and the other sandhi forms. As a result, the between-word pitch distance depended mostly on the SC tonal categories of the first syllables.

3.3.4 Effects of JM tonal merging

Tonal merging in JM generally reduces between-word pitch distance. In the final speaker-wise models, the main effect of merge was significant in all the models, $F \in [31.37, 2870.00]$, $p < 0.05$. Potential merging in JM reduced between-word pitch distance. This effect was more robust for the merging of tonal combinations.

However, Merge also showed complex interactions with the SC tonal relations. The two-way interaction of the tonal relation on the first SC syllables and Merge was significant in 16 of the models, $F \in [3.38, 30.22]$, $p < 0.05$, insignificant in 2 of the models, $F \in [1.84, 2.06]$, $p > 0.05$, and removed in 24 of the models. The two-way interaction of the tonal relation on the second SC syllable and Merge was significant in 26 of the models, $F \in [4.80, 113.40]$, $p < 0.05$, insignificant in 1 of the models (speaker35), $F = 0.87$, $p > 0.05$, and removed in 15 of the models. The three-way interaction of the tonal relations on the first and second SC syllables and Merge was significant in 8 of the models, $F \in [11.01, 66.96]$, $p < 0.05$ and removed in 34 of the models. The post-hoc tests showed that the effect of merge is more robust for word pairs that do not share any tonal category in SC, in that the word pairs with both syllables undergoing tonal merging had smaller between-word distances and the word pairs with merging of tonal combinations had even smaller distance. However, when the two words shared tonal categories in SC, the effect of Merge was disrupted and only the merging of tonal combinations guaranteed the smallest between-word pitch distance for every speaker.

3.3.5 Effect of word frequency

The effect of word frequency was generally inconsistent across speakers, although there was weak evidence supporting that two words from the same word frequency groups tend to have relatively smaller pitch distance.

In the final speaker-wise models, the main effect of word frequency was significant in 26 of the models, $F \in [3.59, 55.91]$, $p < 0.05$ and insignificant in 16 of the models, $F \in [0.02, 2.83]$, $p > 0.05$. The direction of word frequency effects was not consistent across speakers, except that in most models the between-word pitch distances were relatively smaller when both words were high frequency words.

The interaction from word frequency was unclear. The two-way interaction of the tonal relation on the first SC syllables and word frequency was significant in 31 of the models, $F \in [4.65, 71.62]$, $p < 0.05$ and insignificant in 10 of the models, $F \in [0.43, 2.70]$, $p > 0.05$, except that the term was removed for Speaker09. The two-way interaction of the tonal relation on the second SC syllable and word frequency was significant in 20 of the models, $F \in [3.32, 85.86]$, $p < 0.05$ and insignificant in 21 of the models, $F \in [0.07, 2.94]$, $p > 0.05$, except that the term was removed for Speaker25. The two-way interaction of Merge and word frequency was significant in all of the models, $F \in [3.87, 44.69]$, $p < 0.05$, except that the term was removed in 3 models. The two-way interaction of neutral and word frequency was significant in30 of the models, $F \in [3.58, 91.45]$, $p < 0.05$ but insignificant in 12 of the models, $F \in [0.09, 1.95]$, $p > 0.05$. The two-way interaction of between-dialect identity and word frequency was significant in 31 of the models, $F \in [2.52, 37.15]$, $p < 0.05$ and insignificant in 10 of the models, $F \in [0.28, 2.02]$, $p > 0.05$, except that the term was removed for Speaker05. The three-way interaction of the tonal relation on the first SC syllables, Merge, and word frequency was significant in only one of the models $F = 13.10$, $p < 0.05$ but removed in all the other models. The three-way interaction of the tonal relation on the first SC syllables, between-dialect identity, and word frequency was significant in 22 of the models, $F \in [3.03, 24.58]$, $p < 0.05$, but removed in 20 of the models. The three-way interaction of the tonal relation on the first SC syllables, neutral, and word frequency was significant in 24 of the models, $F \in [3.38, 74.36]$, $p < 0.05$, but removed in 18 of the models. The three-way interaction of the tonal relation on the second SC syllable, merge, and word frequency was significant in 6 of the models, $F \in [4.11, 11.22]$, $p < 0.05$ but removed in 36 of the models. The three-way interaction of the tonal relation on the second SC syllable, between-dialect identity, and word frequency was significant in 25 of the models, $F \in [2.38, 17.14]$, $p < 0.05$ but removed in 17 of the models. The three-way interaction of the tonal relation on the second SC syllable, neutral, and word frequency was significant in 29 of the models, $F \in [3.24, 50.36]$, $p < 0.05$, but removed in 13 of the models. The three-way interaction of the tonal relations on the first and second SC syllables and word frequency was significant in 16 of the models, $F \in [3.13, 67.38]$, $p < 0.05$, but removed in 26 of the models. The three-way interaction of neutral tone, between-dialect identity, and word frequency was significant in 11 of the models, $F \in [3.96, 22.79]$, $p < 0.05$, but removed in 31 of the models. The directions of these interactions were inconsistent across speakers except that in most models the effect of sharing tones on the first SC syllables increased when both words were low frequency words.

3.3.6 Effects of the speakers' age, cognitive, and socio-linguistic backgrounds

Individual backgrounds influence the way systematic correspondence takes effect. In this session we start with the predictor age and investigated how the speakers' cognitive and socio-linguistic backgrounds exert age-mediated and age-unmediated influences on systematic correspondence.

Individual models. Before starting investigating the age-independent effects, we first included each aspect of the individual backgrounds separately in smaller models and investigated their mediated and unmediated effects together.

First we started with the age effect. When age was included alone with the linguistic predictors into the model, as shown in Figure 8, a younger age generally increased the between-word pitch distance between two JM words and enhanced the effect of the tonal relations in SC. The main effect of age and the interaction of age with the SC tonal relations were significant when neither of the words was identical to its SC counterpart, F_main (40.00) = 5.23, $p < 0.05$; F_interaction (119.97) = 12.03, $p < 0.05$; and when one word was identical to its SC counterpart, F_main (40.00) = 18.43, $p < 0.05$; F_interaction (119.97) = 5.60, $p < 0.05$, but insignificant when both of the two words were identical to their counterpart in SC, F_main (37.53) = 0.06, n.s.; F_interaction (116.89) = 0.65, n.s. Thus the age effect only existed when at least one word was purely in JM. The proportions of variance accounted for by the final models (R^2) with age are $R\_neither^2 = 0.96$, $R\_one^2 = 0.79$, $R\_both^2 = 0.89$.

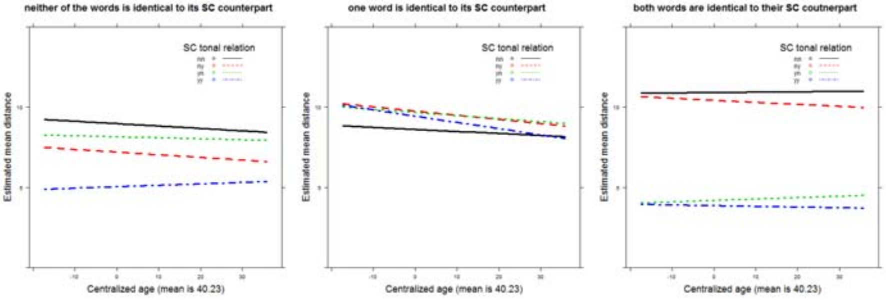


Figure 8. *The interaction of the age and SC tonal relation [neither (nn), only the first (yn), only the* 2^{nd} *(ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words. The estimates were split into three plots according to between dialect identity [neither (left), one (middle), both (right) of the two words is/are identical to its/their counterpart(s) in SC].*

When auditory working memory was included alone with the linguistic predictors into the model, we found that a better auditory working memory enhanced the effect of the SC tonal relations. The main effect of auditory working memory was only significant when one word was identical to its SC counterpart, F (38.00) = 9.95, $p < 0.05$, but insignificant when neither or both of the words was/were identical to its/their SC counterpart(s), F_neither (38.00) = 2.66, n.s.; F_both (35.59) = 0.06, n.s. The interaction of auditory working memory and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (113.97) = 13.88, $p < 0.05$; F_one (113.97) = 4.70, $p < 0.05$, but insignificant when both of the two words were identical to their SC counterparts, F_both (110.96) = 0.48, n.s. The direction of the main effect of auditory working memory when one word was identical to its SC counterpart indicates that a better auditory working memory generally increased the between-word pitch distance between a JM form and a common form. Also it is clear that speakers who have better auditory working memories showed greater differences of between-word pitch distance across different levels of SC tonal relations when neither of the two words was identical to its SC counterpart. The proportions of variance accounted for by the final models (R^2) with auditory working memory are R_neither^2 = 0.96, R_one^2 = 0.78, R_both^2 = 0.88.

When tonal awareness was included alone with the linguistic predictors into the model, we found that a better tonal awareness enhanced the effect of the SC tonal relations. The main effect of SC tonal awareness was only marginally significant when one word was identical to its SC counterpart, F (38.00) = 4.08, p = 0.05, and insignificant when neither or both of the words was/were identical to its/their SC counterpart(s), F_neither (38.00) = 0.84, n.s.; F_both (35.84) = 0.29, n.s. The interaction of SC tonal awareness and SC tonal relations was significant only when neither of the words was identical to its SC counterpart, F_neither (113.97) = 4.03, p < 0.05, but insignificant when neither or both of the words was/were identical to its/their SC counterpart(s), F_one (113.98) = 0.87, n.s.; F_both (111.22) = 0.69, n.s. Similarly, the main effect of JM tonal awareness was only significant when one word was identical to its SC counterpart, F (38.00) = 4.43, p < 0.05, but insignificant when neither or both of the words was/were identical to its/their SC counterpart(s), F_neither (38.00) = 2.35, n.s.; F_both (35.87)= 0.51, n.s. The interaction of JM tonal awareness and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (113.97) = 3.70, p < 0.05; F_one (113.98) = 2.44, p < 0.05, but insignificant when both of the two words were identical to their SC counterparts, F_both (111.26) = 0.15, n.s. The direction of the main effect of tonal awareness when one word was identical to its SC counterpart indicates that a better tonal awareness also generally increased the between-word pitch distance between a JM form and a common form. Also it is clear that speakers who enjoyed better tonal awareness showed greater differences of between-word pitch distance across different levels of SC tonal relations when neither of the two words was identical to its SC counterpart. The proportions of variance accounted for by the final models (R^2) with JM tonal awareness are R_neither^2 = 0.95, R_one^2 = 0.77, R_both^2 = 0.88.

However, digit naming speed showed no significant main effect with any level of the between-dialect identity, F_neither (40.00) = 0.04, n.s., F_one (40.00) = 0.80, n.s., F_both (37.29) = 0.35, n.s., and no significant interaction with the SC tonal relations, F_neither (119.97) = 2.21, n.s., F_one (119.98) = 0.40, n.s., F_both (116.66) = 0.33, n.s. The proportions of variance accounted for by the final models (R^2) with digit naming speed are R_neither^2 = 0.94, R_one^2 = 0.76, R_both^2 = 0.88.

Also, with the pitch range difference across speakers normalized, gender also showed no significant main effect, F_neither (40.00) = 0.06, n.s.; F_one (40.00) = 0.09, n.s.; F_both (38.00) = 0.12, n.s., no significant interaction with the SC tonal relations, F_neither (119.97) = 0.67, n.s.; F_one (119.98) = 0.22, n.s.; F_both (117.40) = 0.19, n.s., and no significant interaction with age, F_neither (35.00) = 1.55, n.s.; F_one (35.00) = 1.57, n.s.; F_both (33.06) = 1.42, n.s. Thus there is no evidence in support of any gender-affected variant choice in JM.

As for the other socio-linguistic predictors, when JM absolute or relative frequency was included alone with the linguistic predictors into the model, we found that a lower frequency in JM enhanced the effects of the SC tonal relations. The main effects of JM relative and absolute frequencies were significant only when one of the words was identical to its SC counterpart, F_relative (40.00) = 22.94, p < 0.05, F_absolute (40.00) = 12.30, p < 0.05, but insignificant when neither or both of the words was/were identical to its/their counterpart in SC, F_neither_relative (40.00) = 2.87, n.s., F_both_relative (38.15) = 0.35, n.s., F_neither_absolute (40) = 1.26, n.s., F_both_absolute (38.27) = 0.12, n.s. The interaction of JM relative frequency with SC tonal relations was significant for all the levels of between-dialect identity, F_neither (119.96) = 27.36, $p < 0.05$, F_one (119.98) = 5.42, $p < 0.05$, F_both (117.46) = 3.51, $p < 0.05$. The interaction of JM absolute frequency and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (119.97) = 16.58, $p < 0.05$, F_one (119.98) = 2.97, $p < 0.05$, but insignificant when both of the two words were identical to their counterpart in SC, F_both (117.60) = 2.10, n.s. The direction of the main effect of JM absolute and relative frequencies when one word was identical to its SC counterpart indicates that the increased usage of JM generally reduces the between-word pitch distance between JM and SC forms. Also it is clear that speakers who use JM relatively or absolutely more often showed smaller differences of between-word pitch distance across different levels of SC tonal relations when neither of the two words was identical to its SC counterpart. The proportions of variance accounted for by the final models (R^{2}) with relative frequency are R_neither^{2} = 0.97, R_one^{2} = 0.79, R_both^{2} = 0.90. The proportions of variance accounted for by the final models (R^{2}) with absolute frequency are R_neither^{2} = 0.96, R_one^{2} = 0.78, R_both^{2} = 0.89.

Similarly, when JM absolute or relative proficiency was included alone with the linguistic predictors into the model, we found that a lower proficiency in JM enhanced the effect of the SC tonal relations. The main effect of JM relative proficiency was significant only when one word was identical to its SC counterpart, F (40.00) = 13.20, $p < 0.05$, but insignificant when neither or both of the words was/were identical to its/their SC counterpart(s), F_neither (40.00) = 1.60, n.s.; F_both (37.25) = 0.17, n.s. The proportions of variance accounted for by the final models (R^{2}) with relative proficiency are R_neither^{2} = 0.95, R_one^{2} = 0.77, R_both^{2} = 0.89. The interaction of JM relative proficiency and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (119.97) = 15.38, $p < 0.05$; F_one (119.98) = 3.53, $p < 0.05$, but insignificant when both of the two words were identical to their SC counterpart, F_both (116.58) = 1.59, n.s. The effect of JM absolute proficiency was similar in that the interaction of JM absolute proficiency and SC tonal relations was significant when neither of the words was identical to its SC counterpart, F_neither (119.97) = 3.29, $p < 0.05$, but insignificant when one or both of the two words were identical to their SC counterparts, F_one (119.98) = 1.40, n.s., F_both (116.83) = 0.32, n.s. However, the main effect of JM absolute proficiency was insignificant for all the levels of between-dialect identity, F_neither (40.00) = 0.01, n.s.; F_one (40.00) = 1.95, n.s.; F_both (37.46) = 0.02, n.s. The direction of the main effect of JM relative proficiency when one word was identical to its SC counterpart indicates that the increased dominance of JM generally reduced the between-word pitch distance between JM and SC forms. Similar to what was found for JM frequencies, speakers who are more proficient in JM showed smaller differences of between-word pitch distance across different levels of SC tonal relations when neither of the two words was identical to its SC counterpart. The proportions of variance accounted for by the final models (R^2) with absolute proficiency are $R\_neither^2 = 0.96$, $R\_one^2 = 0.78$， $R\_both^2 = 0.89$.

Excluding the only speaker who did not receive formal education and including education level or language of literacy education alone with the linguistic predictors into the model, the main effect of education level was only significant when one word was identical to its SC counterpart, F (37.00) = 4.47, $p < 0.05$, but insignificant when neither or both of the words were identical to their SC counterpart(s), F_neither (37.00) = 0.00, n.s.; F_both (34.26) = 1.27, n.s. The interaction of education level and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (110.97) = 9.14, $p < 0.05$; F_one (110.98) = 4.75, $p < 0.05$, but insignificant when both of the two words were identical to their SC counterpart, F_both (107.61) = 0.36, n.s. Similarly, the main effect of the language of literacy education was only significant when one word was identical to its SC counterpart, F (36.00) = 11.16, $p < 0.05$, but insignificant when neither or both of the words were identical to their SC counterpart, F_neither (36.00) = 0.59, n.s.; F_both (33.56) = 0.05, n.s. The interaction of the language of literacy education and SC tonal relations was significant when neither or one of the words was identical to its SC counterpart, F_neither (107.97) = 7.38, $p < 0.05$; F_one (107.98) = 3.13, $p < 0.05$, but insignificant when both of the two words were identical to their SC counterparts, F_both (104.90) = 1.14, n.s. When one of the words was identical to its SC counterpart, a higher education or a literacy education in SC increased the between-word pitch distance and changed the way SC tonal relations influenced acoustic distance. When neither of the words was identical to its SC counterpart, a higher education or a literacy education in SC enhanced the effect of the SC tonal relations. When both of the words were identical to their SC counterparts, the education level and literacy education had no effect. The proportions of variance accounted for by the final models (R^2) with education are $R\_neither^2 = 0.95$, $R\_one^2 = 0.78$， $R\_both^2 = 0.88$. The proportions of variance accounted for by the final models (R^2) with literacy education are $R\_neither^2 = 0.96$, $R\_one^2 = 0.79$， $R\_both^2 = 0.89$.

Taken together, any one of these speaker-related variables has been proven useful when included alone in the modeling. Systematic correspondence has a stronger effect with younger ages, better tonal awareness, larger auditory working memories, literacy educations in SC, higher education levels, and lower frequencies and proficiencies of JM usage, when neither of the words was identical to its SC counterpart. Such backgrounds also increased the between word pitch distance when only one of the JM words was produced identical to its SC counterpart. However, when both of the JM words were produced identically to their SC counterparts, these individual background predictors had little effects.

However, which individual backgrounds have age-independent effects? We addressed this question by including these predictors (after residualization) together with age to see to what extent they still played a significant role in accounting for the between-word distances.

Residualized models. We first analyzed how the other interval predictors of individual backgrounds are related to age and between each other. As expected, complex multi-collinearity exists across the individual backgrounds predictors. Taking all the individual backgrounds together, the Kappa value was 7.77, indicating that the collinearity problem was not very serious in general (Kappa smaller than 10 indicates reasonable collinearity). However, pair-wise Pearson correlations showed that most of the other predictors strongly correlated with age (absolute r ranges from 0.38 to 0.73). The other predictors also showed strong correlations between each other (absolute r > 0.27), except that the digit naming speed seemed to be more independent (absolute r ranged between 0.04 and 0.38). Moreover, the education level and language of literacy education also co-varied with age, in the way that younger speakers tend to receive a higher education and receive the literacy education in SC. Only gender was independent of age.

Accordingly, we residualized all the other interval predictors against age. We also assumed that the proficiency effects are mediated by frequency effects and the effects of JM tonal awareness are mediated by the effects of SC tonal awareness. Thus the JM proficiency was residualized against the combination of JM frequency and age and the tonal awareness of JM was residualized against the combination of SC tonal awareness and age. Instead of the relative frequency and proficiency, the absolute ones were kept because they had smaller correlations with age. The initial model included Speaker as the random intercept and the following terms as fixed effects: SC tonal relations, standardized age, education level, language of literacy education, all the possible two-, three-, and four-way interactions of the above mentioned terms, residualized JM absolute proficiency, residualized JM absolute frequency, residualized JM auditory working memory, residualized JM tonal awareness, residualized SC tonal awareness, and all the possible two- and three-way interactions of each residualized predictor with SC tonal relations and education level.

After statistically removing the collinearity, we managed to test all the individual backgrounds in one bigger model. The result showed that only the lower frequency of JM still had an unmediated effect of strengthening the effect of systematic correspondence.

When neither of the two words was identical to its SC counterpart, the main effects of SC tonal relations, F (86.97) = 201.91, p < 0.05, and education level, F (27.00) = 5.53, p < 0.05 were significant. The main effect of residualized JM frequency, F (27.00) = 0.14, n.s., was insignificant in the final model and the main effect of residualized JM proficiency, F (27.00) = 3.71, p = 0.06, was only marginally significant. All the terms including the residualized auditory working memory and residualized tonal awareness were removed in the final model. The two-way interaction of SC tonal relations and residualized JM frequency, F (86.97) = 3.94, p < 0.05 and the two-way interaction of education level and residualized JM proficiency, F (27.00) = 7.47, p < 0.05 were significant. The tree-way interaction of the SC tonal relations, age, and the language of literacy education was significant, F (86.97) = 2.24, p < 0.05. The three-way interaction of the SC tonal relations, education level, and the language of literacy education was also significant, F (86.97) = 4.39, p < 0.05. Similar to the results from the other types of models, sharing tones in SC reduced the JM between-word pitch distance under the present condition. As shown in Figure 9, after removing the mediated effect from age, a higher JM frequency reduced the differences induced by the SC tonal relations and the between-word pitch distance is increased with a higher level of education. The interaction of age and the SC tonal relations was modulated by the type of literacy education. If the speaker received literacy education in JM, a younger age reduced the differences induced by the SC tonal relations. However, if the speaker received literacy education in SC, a younger age enhanced the differences induced by the SC tonal relations. If the speaker received literacy education in a mixed condition, age did not have much effect. The proportion of variance accounted for by this final models (R^2) is $R\_neither^2 = 0.98$.

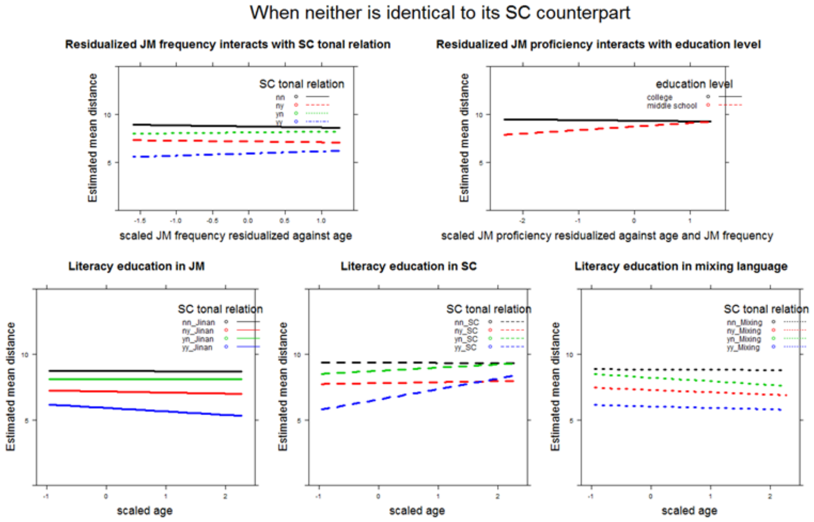


Figure 9. *(1) Above-left: the interaction of the speaker's residualized JM frequency and their interaction with SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words. (2) Above-right: the interaction of the speaker's residualized JM proficiency and the speaker's education level (middle school, college) on the estimated between-word pitch distance of JM words. (3) Below: the interaction of age, literacy education, and SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words. The estimates were split into three plots according to the language of literacy education.*

When only one of the two words was identical to its SC counterpart, only the main effects of SC tonal relations, $F(107.98) = 25.74$, $p < 0.05$, the main effect of the language of literacy education, $F(36.00) = 11.16$, $p < 0.05$, and their interaction, $F(107.98) = 3.13$, $p < 0.05$ were significant. All the other terms were insignificant and removed in the final model. As shown in Figure 10, sharing tonal categories in SC did not reduce but increased between-word pitch distance, because in this case we are comparing corresponding tonal categories in JM and SC. Receiving literacy education in SC increased the between-word pitch distance between a JM form and a common form but receiving literacy education in JM reduced such pitch distance.

The literacy education in SC seems to tear the corresponding tonal categories from the two dialects apart and the literacy education in JM seems to make the corresponding tonal categories acoustically more similar. The proportion of variance accounted for by this final models (R^2) is $R\_one^2 = 0.79$.

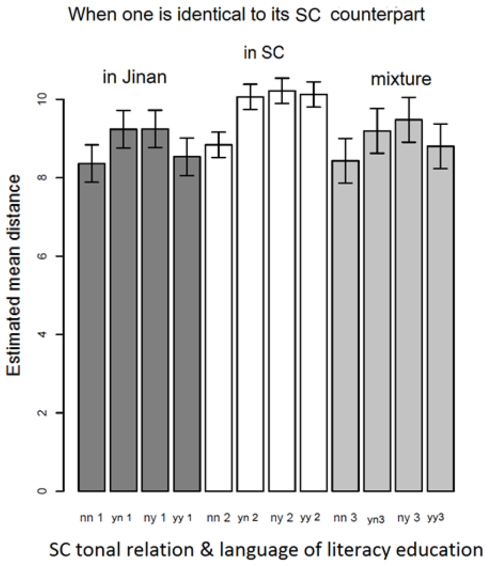


Figure 10. *The interaction of the language of literacy education and tonal relation in SC [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s)] are from the same tonal category] on the estimated between-word pitch distance of JM words.*

When both of the words were identical to their counterpart in SC, only the main effect of SC tonal relations was significant, F (3) = 245.87, $p < 0.05$ in the final model (the random intercept of Speaker was removed in model trimming). All other predictors and interactions were kept but insignificant. This effect was consistent with the results of the speaker-wise models. Sharing tonal categories in SC reduced the JM between-word pitch distance. The proportion of variance accounted for by this final models (R^2) is $R\_both^2 = 0.92$.

Stratified models. The nominal predictors of socio-linguistic backgrounds (except gender) also co-varied with age. Younger speakers tend to receive a higher education and receive the literacy education in SC. This type of collinearity was not considered in the model mentioned above. We further stratified the data, and fitted separate models on the subsets of data to see whether the effects found in the bigger model persist within each group of speakers. (The scale predictors were normalized and residualized again for each subset of data).

Although all six possible combinations of language of literacy education and education level (middle-school-JM, college-JM, middle-school-SC, college-SC, middle school-mix, and college-mix) existed in the data, only two subsets of data, namely middle school-JM and college-SC, had enough data points for fitting separate models. Here we report and discuss the results of these two sets of models. The effects found in the bigger model persisted and more unmediated effects emerged in the middle-school-JM group.

First we studied the speakers who achieved middle school level education and received the literacy education in JM. As shown in Figure 11, they were mostly middle aged. When neither of the two words was identical to its SC counterpart, the main effects of SC tonal relations was significant, F (17.99) = 248.09, $p < 0.05$. Although the main effect of residualized absolute proficiency of JM was insignificant, its interaction with SC tonal relations was significant, F (17.99) = 3.61, $p < 0.05$. The main effects of age, residualized JM absolute frequency, normalized JM auditory working memory, and normalized JM tonal awareness were kept in the final model but insignificant. All the other terms of interaction were removed. When residualized relative frequency and proficiency of JM were used instead of the absolute ones in the modeling, the effect of residualized JM proficiency became insignificant, and the interactions from normalized age, F (14.99) = 3.29, $p < 0.05$, and residualized auditory working memory, F (14.99) = 5.14, $p < 0.05$, became significant. The increase of age and residualized auditory working memory enhanced the differences induced by the SC tonal relations. This suggests that the effect of systematic correspondence is stronger on the oldest speakers and the speakers who have better auditory working memories. The advantage of the oldest speakers was also found in the model depicted in Figure 9 and may be due to the philology training the oldest speakers received. As shown in Figure 11, the residualized JM absolute proficiency enhanced the differences induced by the SC tonal relations, which is in the opposite of the direction of the general effect of JM proficiencies. This suggests that after removing the influences from the times and JM frequency, middle-school-JM speakers with higher JM proficiencies show more respect to the systematic correspondence between their two dialects. The proportions of variance accounted for by these final models (R^2) by the middle-school-JM group are $R\_neither^2 = 0.98$, $R\_one^2 = 0.96$， $R\_both^2 = 1.00$.

For this middle-school-JM group, when one of the two words was identical to its SC counterpart, the main effects of SC tonal relations was significant, F (14.99) = 33.71, $p < 0.05$. Although the main effect of residualized JM absolute proficiency was insignificant, its interaction with SC tonal relations was significant, F (14.99) = 8.33, $p < 0.05$. The interaction between SC tonal relations and SC tonal awareness was also significant, F (14.99) = 6.10, $p < 0.05$. For word pairs sharing tones on the second SC syllables, the between-word pitch distance increased with the increase of residualized JM absolute proficiency. For word pairs sharing tones on the second SC syllables but not on the first SC syllables, the residualized JM absolute proficiency increased with the increase of tonal awareness. These effects reflect the pitch differences between the corresponding SC and JM tonal categories. The results indicate that speakers with higher JM proficiency and better tonal awareness maintain greater pitch differences between the corresponding SC and JM tonal categories.

For this middle school-JM group, when both of the two words were identical to their SC counterparts, the random intercept of speaker was removed in the final model and here we report the results of the full model. The main effects of SC tonal relations were significant, F (3.99) = 521.80, $p < 0.05$. The main effect of residualized JM frequency was significant in the F-statistics, F (3.99) = 8.00, $p < 0.05$, but insignificant in the t-statistics. The interaction between SC tonal relations and SC tonal awareness was insignificant in the F-statistics, but sharing the tonal category on the second syllable but not the first syllable significantly interacted with SC tonal awareness in the t-statistics, t (3.99) = 3.30, $p < 0.05$, in that a higher SC tonal awareness increased the between-word pitch distance. This result indicates that the speakers with a better SC tonal awareness focus more on the tonal contrasts on the first SC syllables.

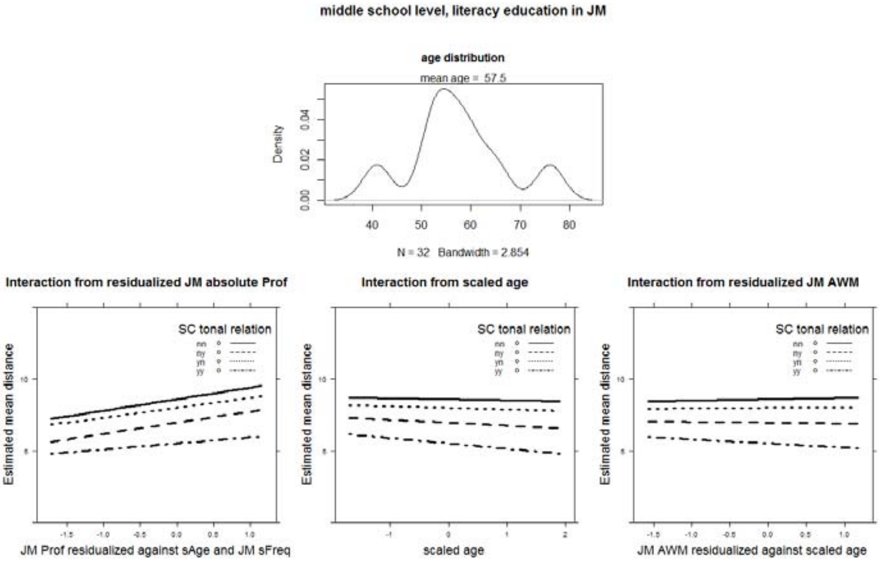


Figure 11. *The models were built for the middle-school-JM group when neither of the two words was identical to its SC counterpart. (1) Above: the age distribution. (2) Below left: the interaction of the speaker's residualized JM absolute proficiency and its interaction with SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words. (3) Below middle: the interaction of the speaker's scaled age and its interaction with SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words. (4) Below right: the interaction of the speaker's residualized auditory working memory and its interaction with SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words.*

Second we studied the speakers who achieved college level education and received the literacy education in SC. As shown in Figure 12, they were mostly in their twenties, younger than the middle school-JM group. When neither of the two words was identical to its SC counterpart, the main effects of SC tonal relations was significant, F (53.99) = 529.45, $p < 0.05$. Although the main effect of residualized absolute frequency of JM was insignificant, its interaction with SC tonal relations was significant, F (53.99) = 3.04, $p < 0.05$. All the other terms were insignificant and removed. We found the same result when residualized relative frequency and proficiency of JM were used instead of the absolute ones in the modeling. As shown in Figure 12, after removing the age effect a higher JM frequency reduced the differences induced by the SC tonal relations for these speakers. Thus the effect of JM frequency on systematic correspondence is not just mediated by the age effect. Within this young group of bilinguals, those who use JM more frequently are less affected by the systematic correspondence. The proportions of variance accounted for by these final models by the college-SC group (R^2) are $R\_neither^2 = 0.96$, $R\_one^2 = 0.68$，$R\_both^2 = 0.87$.

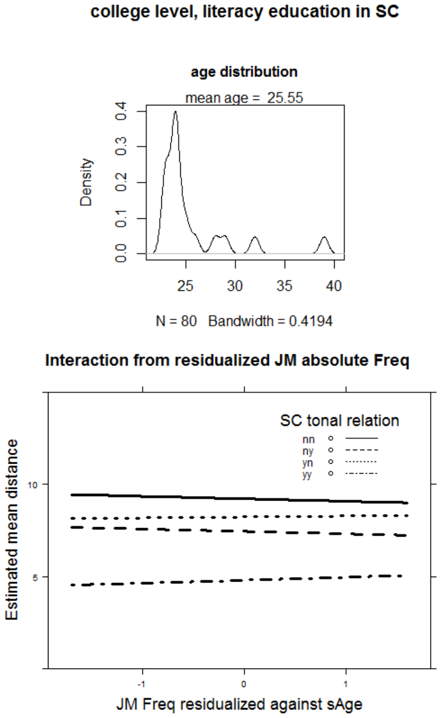


Figure 12. *The models were built for the college-SC group when neither of the two words was identical to its SC counterpart. (1) Above: the age distribution. (2) Below: the interaction of the speaker's residualized JM absolute frequency and its interaction with SC tonal relation [neither (nn), only the first (yn), only the 2nd (ny), both (yy) syllable(s) are from the same tonal category] on the estimated between-word pitch distance of JM words.*

For this college-SC group, when one of the two words was identical to its SC counterpart, only the main effects of SC tonal relations were significant, F (56.99) = 21.86, $p < 0.05$. When both of the two words were identical to their SC counterparts, the random intercept of speaker was removed in the final model and in the full model also only the main effects of SC tonal relations were significant, F (37.70) = 123.02, $p < 0.05$.

Taken together, age effect interacts and covaries with the language of literacy education. Considering all the speakers in this corpus, when the speaker received literacy education in SC, a younger age strengthens the effect of systematic correspondence. However, when the speaker received literacy education in JM, a younger age reduces the effect of systematic correspondence. One explanation is that the oldest people (who received their literacy education in JM) received some training in philology and the youngest people (who received their literacy education in SC) received better education in pinyin. These interactions are only true within the JM system. When considering a pair of JM words that are both realized almost identically to their SC counterparts, only the SC tonal relations is taking charge.

However, most of the participants are either middle-aged people who achieved middle school education and received their literacy educations in JM or young people who achieved college education and received their literacy educations in SC. This demographic distribution is shaped by the social progress in the past 30 years. Analyzing these two main groups of bilinguals separately, more cognitive and sociolinguistic effects emerged with the older group. After statistically removing the mediated effects, we found that the older bilinguals with higher JM proficiencies and better auditory working memories show more respect to the systematic correspondence. Also the older bilinguals with higher JM proficiency and better tonal awareness are better at distinguishing SC and JM tonal categories. We did not find such effects on the younger group. Instead we only found an unmediated effect of JM absolute frequency on these younger bilinguals. Those who use JM more often are less influenced by systematic correspondence.

3.4 General discussion

3.4.1 Main findings

First, we have shown that SC tonal relations, between-dialect identity, and their interactions are significant in predicting the between-word pitch distance in JM. It is not surprising that the present results are in line with the basic principle of historical linguistics, namely that phonemes correspond systematically for related words between two related languages/dialects (Meillet & Ford, 1967). We deduced from this well-known notion that the tonal realizations of JM words could be predicted from the tonal categories of their SC counterparts (see Table 1). We also proved this implication via statistical modeling of JM between-word pitch distance.

There is always a certain proportion of words that do not fit the systematic correspondence rule. These are usually treated as exceptions when historical linguists talk about the rules. However, these ‘exceptions’ need to be handled if we want to apply the systematic correspondence rules in engineering. On the other hand, these ‘exceptions’ may reflect another potentially useful type of relationship between the two languages/dialects, namely phonological similarity. JM is in close and long term contact with SC and most JM speakers are bilinguals with SC as L2. As a result, a large number of the JM ‘exception’ words are almost identical to their SC counterparts, which are sometimes, though neither necessarily nor sufficiently, a result of borrowing (see 1.3 and 1.5). In the present study, we have figured out how this mechanism interacts with the effects from systematic correspondence in predicting JM between-word pitch distance.

Second, neutral tones disrupt systematic correspondence. Neutral tones disrupted the effect of the SC tonal relation on the second syllable. The effect of the SC tonal relation on the second syllable was also disrupted by neutral tones when neither of the two words was identical to its SC counterpart. Only when both words were identical to their SC counterparts, neutral tones instead enhanced the effect of the tonal relations on the first SC syllables. These findings have linguistic basis. Neutral tones, whether in JM or SC, are realized differently depending on the previous tonal categories. Thus simply sharing neutral tones does not guarantee a smaller between-word pitch distance. Nevertheless, neutral tone sandhi works differently in SC and JM. JM morphemes followed by neutral tones carry different tonal variants from their citation forms and other sandhi forms. However, SC morphemes followed by neutral tones realize similarly to the corresponding citation forms. These findings taken together suggest that, in order to improve the predicting power of systematic correspondence when converting a JM pronunciation dictionary from a SC one, it may be better to distinguish subtypes of neutral tones by their preceding tones and also include whether the current syllable is followed by a neutral tone.

Third, we found out how cognitive and sociolinguistic backgrounds affect the way systematic correspondence works, where we differentiated the age-mediated and age-independent effects of these individual backgrounds.

We started with the general (mediated and unmediated together) effects of individual backgrounds. Including each aspect of individual backgrounds separately into the model, we found that the younger, less frequent, less proficient speakers of JM, who received higher education, received their literacy educations in SC, and have better working memories, were generally more sensitive to the SC tonal relations and produced the JM and SC corresponding tonal categories with greater differences. These findings reflect the change of the JM tonal system across different generations: the tonal relations across JM words are increasingly aligned with the SC tonal relations but the acoustic realizations of the JM tonal categories are less and less similar to their corresponding tonal categories in SC. With the changes in society, the systematic correspondence mechanism has been strengthened and the surface tonal similarity has been reduced.

However, after the age-mediated effects were statistically removed from the individual backgrounds and different speaker-related predictors were modeled together, more complex interactions emerged. First, JM frequency showed age-independent effects. The more a speaker uses JM, the less sensitive he or she is to the SC tonal relations compared with his or her peers. Second, the more highly-educated speakers and the more proficient speakers generally maintain a greater JM between-word pitch distance than their peers. Third, for the speakers who received their literacy educations in JM, age increases their sensitivity to SC tonal relations and the pitch distances between JM and SC corresponding tonal categories increase. However, for the speakers who received their literacy educations in SC, age reduces their sensitivity to SC tonal relations. This discontinuity of age effect may be attributed to the changes in education. The younger speakers received pinyin education and the oldest speakers received traditional philology training. Only the middle-aged speakers suffered from their deprivation of formal primary education. Taken together, with age-mediated effects excluded, the frequency of language usage and the educational experiences influence the bilingual individuals' sensitivity to systematic correspondence and the way they maintain the tonal contrasts.

The effects of cognitive background emerged after we focused on the two representative groups of speakers. The young speakers who received their literacy educations in SC and completed higher education are only sensitive to JM frequency and did not show any age-independent effects from their cognitive backgrounds. Neither did they show any age effects. However, with the middle-aged speakers whose aging has already started, we found significant effects of JM proficiency independent of age and JM frequency. The unmediated effect of JM proficiency may reflect some influence from their language aptitudes. The middle-aged individuals who can maintain relatively higher JM proficiencies usually are also more respectful to the systematic correspondence across the two tonal systems. Similarly, the middle-aged speakers who maintain a better auditory working memory are also more sensitive to SC tonal relations compared with their peers. Individuals are shaped by their experiences when they are young. When cognitive aging happens, different individuals age at different speeds and in different ways, which endow their tonal systems with different colors.

The effects of individual backgrounds are in line with earlier studies of JM. Researchers have been studying the age-related and style-related segmental variations in JM for more than 20 years (Cao, 1991; Z.-Y. Qian, 1997), although they did not include tonal variability in their discussion. Our findings are also in line with the basic principles of sociolinguistics, in that social contexts influence the choice of linguistic variants and speech performances within the linguistic society (Labov, 2006; Weinreich et al., 1968). However, most phonological and phonetic sociolinguistic studies worked on how individual backgrounds affect specific variants, while the current study explored how individual backgrounds affect the general mechanisms which control the alignment of two tonal systems. The cognitive effects are related to previous finding on bilinguals (Bialystok et al., 2008; Bialystok, Poarch, Luo, & Craik, 2014; Signorelli et al., 2011). However, the effect of auditory working memory on the strength of systematic correspondence is relatively new.

We did not focus on the acoustic variability [as shown in Figure 1(a)] which can be handled via speech normalization and changing the parameters in the computational models (Leggetter & Woodland, 1995; Woodland, 2001). We managed to improve the analysis of tonal pattern variability [as shown in Figure 1(b)] by including individual backgrounds. Note that the tonal pattern variability also exists within the same speaker, which cannot be directly handled by grouping the speakers or building different accent models (Huang et al., 2000; Woodland, 2001). The current results suggest that individual backgrounds do not necessarily work in a 'yes-no' way in regulating the selection of tonal patterns. Individual backgrounds may instead affect the relative probability of different tonal patterns of the same word considering how the JM words are related to their SC counterparts.

3.4.2 Limitations

The present study has several limitations. First, the study only involved isolated disyllabic words. Although the majority of modern Chinese words are disyllabic and the uncertainty of disyllabic tonal realizations has partly been explained with individual backgrounds, how disyllabic JM words are realized in connected speech needs further investigation. Researchers have done fruitful studies on contextual tonal realizations and their interactions with sentential prosodies (Chen, 2010; Chen and Gussenhoven, 2008; Xu, 1997, 1999; Xu and Prom-on, 2014; Xu and Wang, 2001). However, how to transfer this knowledge from SC to the other Chinese dialects and how the predictors we investigated work in context are still open questions.

Second, the JM pronunciations were only obtained via a word reading task. Reading biases the speakers towards a more formal and literal style. As a result, the frequency of some variants in our corpus may diverge from the real probability in everyday speech. Moreover, we do not know how some speakers can read the words in JM without being taught to read in JM. Without enough background knowledge on this question, we do not know the exact effect this has on their JM pronunciation.

Third, whether a rendition of a word was produced almost identically to its SC counterpart was manually marked. This information is necessary for incorporating the phonological similarity mechanism in predicting the JM tonal realization. However, systematic correspondence can make predictions in some cases: the JM neutral tone sandhi with high-level as the first citation tone usually results in sandhi forms almost identical to their SC counterparts; thus, we can predict that the SC words with low-rising+neutral tones are more likely to have identical counterparts in JM. Nevertheless, the rest of the cases are either conventional (most speakers do so for this word but other words from the same tonal categories in SC are not treated the same) or incidental (only a few speakers do so). The rule-based between-dialect identity can be easily predicted and the conventional between-dialect identity (borrowing) needs to be marked manually. How to build a recognizer that can handle the incidental between-dialect identity (code-mixing) is still an open question.

Fourth, every part of one pitch contour received the same weight in the calculation of Euclidean distance. Euclidean distance serves as an objective measurement for the difference between two pitch contours. However, human perception is more complex. On the one hand, different parts of pitch contour contribute differently to speech perception (Whalen & Xu, 1992). On the other hand, tonal perception is incremental (Shen, Deutsch, & Rayner, 2013). Tones can be retrieved quite early (Lee, 2000) and different tones have different isolation points (Lai & Zhang, 2008), which means introducing the interaction between tonal category and the weights on different parts of the contour may benefit the modeling. Further studies may consider incorporating this knowledge into the weighting of different parts of the pitch contour.

Fifth, the onset condition was not taken into consideration. Previous studies (Y. Chen, 2011; Howie, 1974; C. X. Xu & Xu, 2003) have shown that pitch contours carried by the rhymes are influenced by the onset condition. For instance, given two word pairs, both involving SC words from the same tonal categories, the pair with the same type of onsets would be more similar in pitch contour than the pair with different types of onsets. This bias would be also applicable to the JM translation equivalents, since they are almost identical to their SC counter parts on the segmental level. This predictor can be taken into consideration in further studies.

Sixth, the middle Chinese tonal system is known but not taken into consideration in this study. The major systematic discrepancy between SC and JM (and also between SC and other Mandarin dialects) comes from the different redistribution of the historical 'Rusheng' (checked tone) category (Collective_work, 1989). This knowledge, although not applicable for other languages, can be potentially beneficial for the modeling of Chinese dialects.

Finally, the methodology used in this study involves two statistical issues. The usage of the distance matrix as the dependent variable for Linear Mixed Modeling introduced an inherent autocorrelation problem. As a result, the model may be over-confident. We carried out post-hoc analysis of model estimates, with the hope that they are more trustworthy for evaluating the effects of the predictors. Another issue is residualization. Recent study suggests that residualization, contra to common recognition in the literature, does not change the result for the residualized variable but the variable which is residualized against (Wurm & Fisicaro, 2014). This indicates that some more attention should be given to the predictor age, against which we did the residualization. However, the interpretation of our findings remains the same. Note that the present study did not include any model with all the original aspects of individual backgrounds fitted together. Hence, neither did we make comparisons between such a model and the residualized models. Different aspects of individual backgrounds were first modeled alone and then integrated together in residualized models. It is because of this integration that the age-independent effects emerged, regardless of whether the models are residualized or not. Finally, of course, other methods are to be further explored in the future to address our research questions.

3.4.3 Applications

Using the systematic correspondence between JM and SC tonal systems, with the between-dialect identity, JM neutral tone and tonal merging, and individual backgrounds under consideration, we have proven that JM lexical tonal realizations can be predicted from the tonal categories of their SC counterparts. This finding has implications for both linguists and engineers. On the one hand, we have shown that individual backgrounds not only affect the selection of specific tonal variants but also contribute to how the general systematic correspondence mechanism takes effect. On the other hand, based on the present findings, further implicational models can more economically transfer SC tonal information in order to build pronunciation dictionaries for the speech recognizers and synthesizers of JM. This approach can probably be generalized to other related northern Chinese dialects and help multi-accent/lingual modeling.

Acknowledgements

We would like to thank Prof. Shengli Cheng, Prof. Xiufang Du, Jia Li, Lulu Zhou, and Dianliu Neighbourhood Committee for the recruitment of participants. We also would like to thank Martijn Wieling for his advice on statistics. J. Wu's work was supported by a PhD Studentship sponsored by Talent and Training China-Netherlands Program. We would like to thank the support to Yiya Chen from the European Research Council (ERC-Starting Grant 206198). The fieldtrip was sponsored by the Leiden University Foundation and Leiden University Centre for Linguistics.

References

Baayen, R. H. (2011). languageR: Data sets and functions with 'Analyzing Linguistic Data: A practical introduction to statistics' R package version 1.4. http://CRAN.R-project.org/package=languageR.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55*(2), 290-313.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity* (Vol. 571): John Wiley & Sons.

Besacier, L., Barnard, E., Karpov, A., & Schultz, T. (2014). Automatic speech recognition for under-resourced languages: A survey. *Speech Communication, 56*(0), 85-100. doi: 10.1016/j.specom.2013.07.008

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. Amsterdam: Glott International. Retrieved from http://www.praat.org

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). *LOF: identifying density-based local outliers.* Paper presented at the ACM Sigmod Record.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729. doi: 10.1371/journal.pone.0010729

Cao, Z. (1991). 'Jinan Fangyan Ruogan Shengmu De Fenbu He Yanbian -- Jinan Fangyan Dingliang Yanjiu Zhiyi' (The distribution and evolution of certain initial consonants in Jinan, quantified study on Jinan I). *Language Study, 2*(004).

Chen, M. (1973). Cross-Dialectal Comparison: A Case Study and Some Theoretical Considerations. *Journal of Chinese Linguistics, 1*(1), 38-63.

Chen, Y., Shen, R., & Schiller, N. O. (2011). *Representation of allophonic tone sandhi variants.* Paper presented at the Proceedings of Psycholinguistics Representation of Tone. Satellite Workshop to ICPhS, Hongkong.

Collective_work (Ed.). (1989). *'Hanyu Fangyin Zihui' (Phonetic Dictionary of Chinese Dialects)*. Beijing: Dept of Chinese Language and Literature, Peking University.

Deza, M. M., & Deza, E. (2009). Euclidean distance *Encyclopedia of distances* (pp. 94): Springer.

Dyen, I. (1963). Why phonetic change is regular. *Language, 39*(4), 631-637. doi: 10.2307/411958

Feng, S. (2001). Cong yunlu kan hanyu ciyu fenliu zhi dajie (Prosodically-determined distinction between word and phrase in Chinese). *Zhongguo Yuwen (Chinese Language)(1)*, 27-37.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103-127. doi: 10.1080/09658219408258940

Hamed, M. B. (2005). Neighbour-nets portray the Chinese dialect continuum and the linguistic legacy of China's demic history. *Proceedings of the Royal Society B: Biological Sciences, 272*(1567), 1015-1022. doi: 10.1098/rspb.2004.3015

Harper, D. (2001). Online etymology dictionary. *Online Etymology Dictionary.* 19, from http://www.etymonline.com/index.php

Huang, C., Chang, E., Zhou, J., & Lee, K.-F. (2000). *Accent modeling based on pronunciation dictionary adaptation for large vocabulary Mandarin speech recognition.* Paper presented at the INTERSPEECH.

Imseng, D., Motlicek, P., Bourlard, H., & Garner, P. N. (2014). Using out-of-language data to improve an under-resourced speech recognizer. *Speech Communication, 56*(0), 142-151. doi: 10.1016/j.specom.2013.01.007

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*(1), 23-62. doi: 10.1016/j.cogpsych.2010.02.002

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://CRAN.R-project.org/package=lmerTest

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics, 34*(4), 500-515. doi: 10.1016/j.wocn.2006.05.002

Leggetter, C. J., & Woodland, P. C. (1995). Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models. *Computer Speech & Language, 9*(2), 171-185. doi: 10.1006/csla.1995.0010

Li, A. (2002). *Chinese prosody and prosodic labeling of spontaneous speech.* Paper presented at the Speech Prosody 2002, International Conference.

Li, Q., & Chen, Y. (2012). *Trisyllabic Tone Sandhi in Tianjin Mandarin.* Paper presented at the Proceedings of the Third International Symposium on Tonal Aspects of Languages.

Li, R. (1988). *The Language Atlas of China*. London: Longman.

Meillet, A., & Ford, G. B. (1967). *The comparative method in historical linguistics*. Paris: H. Champion.

Nixon, J. S., Chen, Y., & Schiller, N. O. (2014). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. *Language, Cognition and Neuroscience*(ahead-of-print), 1-15. doi: 10.1080/23273798.2014.942326

Norman, J. (2003). The Chinese dialects: phonology. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 72-83): Psychology Press.

Peng, S. H. (2000). Lexical versus 'phonological'representations of Mandarin Sandhi tones *Papers in Laboratory Phonology V: Acquisition and the Lexicon (Vol 5)* (pp. 152-167).

Qian, Y., & Soong, F. K.-P. (2012). US8244534 B2.

Qian, Z.-Y. (Ed.). (1997). *'Jinan Fangyan Cidian' (Jinan Dialect Dictionary)*. Nanjing: Jiangsu Education Press.

Qian, Z.-Y., & Zhu, G.-Q. (1998). 'Jinanhua Yindang' (The Sound System of Jinan Dialect): Shanghai Educational Publishing House, Shanghai.

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

Schultz, T. (2002). *Globalphone: a multilingual speech and text database developed at karlsruhe university.* Paper presented at the INTERSPEECH.

Shu, H., Peng, H., & McBride - Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11*(1), 171-181. doi: 10.1111/j.1467-7687.2007.00654.x

Sproat, R., Zheng, F., Gu, L., Jurafsky, D., Shafran, I., Li, J. (2004). Dialectal chinese speech recognition: Final report. doi:10.1.1.133.7189

Tang, C., & van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua, 119*(5), 709-732. doi: 10.1016/j.lingua.2008.10.001

Weinreich, U., Labov, W., & Herzog, M. I. (1968). *Empirical foundations for a theory of language change*: University of Texas Press.

Woodland, P. C. (2001). *Speaker adaptation for continuous density HMMs: A review.* Paper presented at the ISCA Tutorial and Research Workshop (ITRW) on Adaptation Methods for Speech Recognition.

Wu, A., & Jiang, Z. (2000). *Statistically-enhanced new word identification in a rule-based Chinese system.* Paper presented at the Proceedings of the second workshop on Chinese language processing: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 12.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America, 95*(4), 2240-2253. doi: 10.1121/1.408684

Yuan, J. H., & Chen, Y. Y. (2014). 3rd tone sandhi in standard Chinese: A corpus approach. *Journal of Chinese Linguistics, 42*(1), 218-237.

Appendix. Model estimates averaged over the speaker-wise models

Predictors	mean Est.	mean SE	mean df	mean t	mean p
(Intercept)	3.89	0.16	602.35	24.96	0.00
BDidentity one	0.06	0.09	74845.14	2.20	0.09
BDidentity one: WFreq both high	0.01	0.09	74792.72	0.50	0.16
BDidentity one: WFreq both low	-0.01	0.11	74813.59	-0.21	0.15
BDidentity both	1.73	1.05	74771.55	2.01	0.23
BDidentity both: WFreq both high	0.02	0.57	74796.19	0.06	0.27
BDidentity both: WFreq both low	0.13	1.03	74785.24	0.04	0.34
WFreq both high	-0.43	0.13	74884.92	-3.52	0.08
WFreq both low	0.03	0.15	74894.17	0.45	0.20
Merge neither	5.59	0.09	74702.31	60.05	0.00
Merge neither: WFreq both high	0.36	0.13	75418.47	2.78	0.12
Merge neither: WFreq both low	-0.07	0.15	75429.54	-0.49	0.19
Merge 2nd Syl	4.02	0.13	73576.72	31.31	0.00
Merge 2nd Syl: WFreq both high	0.28	0.17	75414.59	1.63	0.22
Merge 2nd Syl:WFreq both low	-0.02	0.18	75423.47	-0.08	0.30
Merge 1st Syl	4.74	0.10	73180.79	46.85	0.00
Merge 1st Syl: WFreq both high	0.34	0.13	75418.58	2.55	0.13
Merge 1st Syl: WFreq both low	-0.21	0.15	75428.50	-1.40	0.24
Merge both	1.55	0.17	72791.94	9.30	0.03
Merge both: WFreq both high	0.13	0.23	75413.42	0.56	0.39
Merge both: WFreq both low	-0.01	0.23	75420.54	-0.03	0.35
NtrlTone yes	0.12	0.04	74035.05	2.81	0.08
NtrlTone yes: BDidentity one	1.42	0.11	74059.00	13.75	0.02
NtrlTone yes: BDidentity one: WFreq both low	-0.31	0.16	76045.24	-2.16	0.04
NtrlTone yes: BDidentity both	0.32	1.25	74153.33	0.66	0.11
NtrlTone yes: BDidentity both: WFreq both high	0.49	1.02	76032.87	0.51	0.47
NtrlTone yes: BDidentity both: WFreq both low	-0.72	1.61	76026.85	-0.50	0.45
NtrlTone yes: WFreq both high	0.06	0.06	74983.80	1.10	0.18
NtrlTone yes:WFreq both low	-0.03	0.06	74987.41	-0.35	0.18
rel.SC1 same	-1.27	0.37	2886.25	-4.31	0.11
rel.SC1 same: BDidentity one	0.36	0.14	73480.24	1.17	0.01
rel.SC1 same: BDidentity one:	-0.17	0.18	73672.32	-0.94	0.24

WFreq both low					
rel.SC1 same: BDidentity both	-5.46	0.94	74592.06	-10.58	0.10
rel.SC1 same: BDidentity both: WFreq both high	-0.28	0.57	73636.59	-0.32	0.19
rel.SC1 same: BDidentity both: WFreq both low	0.05	1.09	73617.88	-0.01	0.27
rel.SC1 same: WFreq both high	-0.01	0.08	74880.07	0.37	0.14
rel.SC1 same: WFreq both low	-0.12	0.08	74883.00	-2.54	0.12
rel.SC1 same: Merge neither	-0.92	0.46	72838.08	-2.20	0.26
rel.SC1 same: Merge neither: WFreq both high	1.95	0.65	78317.57	2.98	0.00
rel.SC1 same: Merge neither: WFreq both low	-2.47	0.70	78314.53	-3.52	0.00
rel.SC1 same: Merge 2nd Syl	-1.65	0.48	72937.14	-3.64	0.09
rel.SC1 same: Merge 2nd Syl: WFreq both high	2.51	0.68	78317.32	3.70	0.00
rel.SC1 same: Merge 2nd Syl: WFreq both low	-3.51	0.72	78315.19	-4.88	0.00
rel.SC1 same: Merge 1st Syl	-1.39	0.47	72869.11	-3.16	0.17
rel.SC1 same: Merge 1st Syl: WFreq both high	1.46	0.69	78318.53	2.13	0.03
rel.SC1 same: Merge 1st Syl: WFreq both low	-2.55	0.75	78322.31	-3.40	0.00
rel.SC1 same: Merge both	1.74	1.55	72825.42	1.21	0.33
rel.SC1 same: Merge both: WFreq both high	-0.55	1.72	78323.32	-0.32	0.75
rel.SC1 same: Merge both: WFreq both low	-3.26	1.14	78312.09	-2.86	0.00
rel.SC1 same: NtrlTone yes	1.65	0.08	73124.22	21.09	0.00
rel.SC1 same: NtrlTone yes: BDidentity one	-3.00	0.19	71622.46	-16.21	0.00
rel.SC1 same: NtrlTone yes: BDidentity both	-6.10	1.47	72334.85	-5.17	0.05
rel.SC1 same: NtrlTone yes: WFreq both high	-0.33	0.11	75338.28	-3.14	0.05
rel.SC1 same: NtrlTone yes: WFreq both low	0.41	0.11	75343.70	3.82	0.06
rel.SC1 same: rel.SC2 same	-1.51	0.75	8280.33	-2.77	0.09
rel.SC1 same: rel.SC2 same: BDidentity one	1.94	0.25	67587.54	7.67	0.02
rel.SC1 same: rel.SC2 same: BDidentity both	2.46	0.73	65456.72	3.39	0.07
rel.SC1 same: rel.SC2 same: WFreq both high	0.17	0.13	75594.58	1.27	0.14

rel.SC1 same: rel.SC2 same: WFreq both low	-0.42	0.13	75594.34	-3.09	0.09
rel.SC1 same: rel.SC2 same: Mergenn	-0.09	1.84	76552.13	-0.03	0.64
rel.SC1 same: rel.SC2 same: Merge 2nd Syl	3.10	2.22	76564.02	1.57	0.22
rel.SC1 same: rel.SC2 same: Merge 1st Syl	2.70	1.87	76577.84	1.61	0.22
rel.SC1 same: rel.SC2 same: Merge both	-0.44	2.72	76533.48	-0.38	0.54
rel.SC1 same: rel.SC2 same: NtrlTone yes	0.56	0.18	68082.55	3.17	0.01
rel.SC2 same	-2.47	0.33	488.72	-7.68	0.00
rel.SC2 same: BDidentity one	0.08	0.14	73928.97	0.12	0.16
rel.SC2 same: BDidentity one: WFreq both high	-0.42	0.16	75341.35	-2.56	0.07
rel.SC2 same: BDidentity one: WFreq both low	0.36	0.20	75379.07	1.79	0.15
rel.SC2 same: BDidentity both	-1.00	0.86	74652.89	-1.89	0.24
rel.SC2 same: BDidentity both: WFreq both high	-0.28	0.53	75378.10	-0.51	0.48
rel.SC2 same: BDidentity both: WFreq both low	0.10	0.99	75208.04	0.08	0.37
rel.SC2 same: WFreq both high	0.15	0.12	74791.89	0.27	0.20
rel.SC2 same: WFreq both low	-0.29	0.12	74793.66	-1.95	0.14
rel.SC2 same: Mergenn	0.30	0.21	73864.77	1.32	0.14
rel.SC2 same: Mergenn: WFreq both high	-0.87	0.40	73006.28	-2.14	0.15
rel.SC2 same: Mergenn: WFreq both low	0.91	0.40	73005.69	2.30	0.04
rel.SC2 same: Merge 2nd Syl	2.52	0.83	73838.07	3.46	0.10
rel.SC2 same: Merge 2nd Syl: WFreq both high	-0.74	0.98	73008.28	-0.75	0.49
rel.SC2 same: Merge 2nd Syl: WFreq both low	1.12	0.70	73005.59	1.46	0.17
rel.SC2 same: Merge 1st Syl	-0.64	0.23	73224.64	-3.19	0.10
rel.SC2 same: Merge 1st Syl: WFreq both high	-0.49	0.41	73008.19	-1.17	0.42
rel.SC2 same: Merge 1st Syl: WFreq both low	0.38	0.40	73007.09	0.97	0.42
rel.SC2 same: Merge both	2.50	1.31	73802.31	2.70	0.20
rel.SC2 same: Merge both: WFreq both high	-0.67	1.59	72997.62	-0.44	0.63
rel.SC2 same: Merge both: WFreq	2.99	1.31	73001.22	2.14	0.12

both low					
rel.SC2 same: NtrlTone yes	2.61	0.09	73647.17	29.40	0.00
rel.SC2 same: NtrlTone yes: BDidentity one	-0.78	0.21	72818.36	-3.61	0.04
rel.SC2 same: NtrlTone yes: BDidentity both	-2.87	1.74	74365.52	-2.33	0.27
rel.SC2 same: NtrlTone yes: WFreq both high	0.01	0.12	74318.13	-0.02	0.04
rel.SC2 same: NtrlTone yes: WFreq both low	0.07	0.13	74326.34	0.46	0.07

* BDidentity: Between-dialect identity [neither (baseline), one, or both]
Merge: Merge [the combination (baseline), neither, 2nd syllable, 1st syllable, or both syllables]
rel.SC1: Tonal relation on the first SC syllables [different (baseline) or same]
rel.SC2: Tonal relation on the second SC syllables [different (baseline) or same]
NtrlTone: Neutral tone [no (baseline) or yes]
WFreq: Word Frequency [different (baseline), both high, or both low]

4 Tonal Similarity Effect: The Role of Tone in the Auditory Lexical Access of Etymologically Related Translation Equivalents

Abstract

Phonological similarity affects the bilingual lexical access of etymologically related translation equivalents (called 'cognates' in earlier studies). Jinan Mandarin (JM) and Standard Chinese (SC) are closely related and share many etymologically related translation equivalents, which are usually orthographically and segmentally identical but vary in tonal similarity. In the present study, we studied SC-JM bilinguals' lexical storage of tone-identical (S+T+) and tone-non-identical (S+T−) translation equivalents, using an auditory lexical decision experiment. S+T+ and S+T− words were presented to SC tonal monolinguals in SC, and to SC-JM tonal bilinguals in both JM and SC. After controlling the possibility of within- and between-dialect repetition priming and other covariates, S+T+ and S+T− items did not show any difference in reaction time. However, the discontinuity of tonal similarity effects and the language-dominance effect support that both types of translation equivalents have dialect-dependent separate representations in the bilingual lexicon, and the retrieval is modulated by the language mode and the bilinguals' attention.

4.1 Introduction

Etymologically-related translation equivalents have a common origin and are similar in sound. They are either inherited from the common ancestor language as cognates or borrowed across languages as loan words. Using 'cognate' to refer to all such words, psycholinguists have found 'cognate facilitation effect' in many different tasks and conditions. 'Cognates' (historically related words) are produced faster than 'non-cognates' in visual word naming (Costa, Caramazza, & Sebastian-Galles, 2000; Hoshino & Kroll, 2008) and picture naming (Costa, Santesteban, & Caño, 2005), and recognized faster in visual lexical decision (Brenders, van Hell, & Dijkstra, 2011; Bultena, Dijkstra, & van Hell, 2012; Dijkstra, Grainger, & Van Heuven, 1999; Duyck, Assche, Drieghe, & Hartsuiker, 2007; Lemhöfer & Dijkstra, 2004), progressive de-masking (Dijkstra et al., 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010), word associations (Van Hell & Dijkstra, 2002), eye-tracked natural reading (Duyck et al., 2007), and self-paced reading (Bultena, Dijkstra, & van Hell, 2013)[7]. Moreover, cognates prime each other like within-language variations (Cristoffanini, Kirsner, & Milech, 1986) and they increase the participants' tendency to switch at the same position as in the prime sentence in the structure priming paradigm (Kootstra, van Hell, & Dijkstra, 2012). Cognate facilitation effect applies to many different language combinations. In a study

---

[7] For L2 learning children, cognate facilitation effects can be interrupted by introducing false-friends in the testing list (Brenders et.al, 2011).

involving different specific bilingual language combinations (Lemhöfer et al., 2008), the cognate status was found to be the only between-language variable that significantly affected the reaction time of L2 word identification, again shortening the reaction times. The cognate facilitation effect has been taken as an important phenomenon which reveals how the bilingual mental lexicon is organized and accessed. It supports the view that the bilingual mental lexicon is integrated and words from different languages are co-activated.

A pair of translation equivalents can be different in several dimensions, such as semantics, orthography, phonology, and relative frequency. Semantic and orthographic similarities have shown facilitatory effects in earlier studies (Dijkstra et al., 1999; Dijkstra et al., 2010). However, the effect of phonological similarity is inconsistent within and across studies (Dijkstra et al., 1999; Dijkstra et al., 2010; Duyck et al., 2007; Lemhöfer & Dijkstra, 2004). Several reasons may be responsible for the unstable effects.

First, the phonological similarity co-varies and interacts with the orthographic similarity. A recent study by Dijkstra et al. (2010) has provided important insights into the interaction of orthography and phonology. By distinguishing orthographically identical and non-identical Dutch-English cognates, they found that the orthographically identical cognates showed not only a large processing advantage compared to the non-identical ones, but also differed in their response to phonological similarity. The identical cognates were subject to facilitation from phonological similarity but the non-identical cognates were comparatively less sensitive to phonological similarity. Dijkstra and colleagues made a claim that pairs of non-identical cognates are stored as two separate lexical nodes and lexical access was slowed down by lateral inhibition. In contrast, pairs of identical cognates are stored as one common lexical node and hence not inhibited but facilitated by phonological co-activation.

Second, etymologically related words can vary along different phonemic dimensions. For instance, some have different vowels and others have different consonants. Nevertheless, considering the contribution of different phonemic dimensions, the specific bilingual language combinations investigated in the previous studies suffer from the relative scarcity of target words. This encouraged us to look for a better test case for the effect of phonological similarity along one specific phonemic dimension.

The bilingualism of Standard Chinese (SC) and Jinan Mandarin (JM) allows us to zoom into one aspect of phonological similarity while keeping other aspects constant. SC and JM are closely related Mandarin dialects. They share many etymologically related translation equivalents which are usually orthographically and segmentally identical but vary in tonal similarity. First, unlike in the previous studies, the etymologically related words are the majority in the vocabulary of these bilinguals. As a result, the majority of translation equivalents sound similar, providing many more test cases. Second, both JM and SC are written with the same logographic Chinese writing system. Thus all JM-SC historically related words are orthographically identical, which avoids the potential confound by orthography. Previous research has benefited from this logographic writing system. For instance, previous studies of speech production have used Chinese to tear apart orthographic and phonological effects (Qingfang Zhang & Weekes, 2009; Zhao, La Heij, &

Schiller, 2012). However, one common logographic writing system in two dialects used by the same bilingual speaker has been little studied. Third, the segmental differences between the historically related words are almost annihilated in the youngest bilinguals. As a result, the two translation equivalents are only different in tone. The tonal similarity between a JM word and its SC translation equivalent determines their phonological similarity. Also, while cognates can be semantically different, the translation equivalents are semantically identical, which is convenient for our experimental control. The case of SC-JM tonal bilingualism allows us to focus on the role of tone in the bilingual lexical representation and access.

As suprasegmental phonemes, lexical tones are both common and special. In production, neither sharing Mandarin tones alone (J. Y. Chen, Chen, & Dell, 2002) nor sharing both surface tones and segments (Y. Chen, Shen, & Schiller, 2011) caused facilitatory priming, unlike the case of segments. However, tonal sharing (tonemes or overt tonal realizations) accompanied with segmental sharing introduced phonological facilitation in two picture-word interference experiments (Nixon, Chen, & Schiller, 2014), similar to the case of segments. In phonological encoding, reaction times in a phoneme monitoring task (Ye & Connine, 1999) and reaction times as well as onset latency of the N200 component in a go/no-go task (Q. Zhang & Damian, 2009) showed that segmental information became available prior to tonal information. However, lexical adaptation seems to work similarly in tones and consonants (McQueen, Cutler, & Norris, 2006; Mitterer, Chen, & Zhou, 2011). The present study can provide further insights for the role of tone in lexical representation.

Using prosodic cues on the lexical level, lexical tones should function similarly to lexical stress in the mental representation of words as abstract lexical frames (Levelt, Roelofs, & Meyer, 1999). Dutch minimal stress pairs did not prime each other compared with neutral controls with two unrelated words (Cutler & Van Donselaar, 2001; Jongenburger, 1996). JM tonal minimal pairs primed each other negatively in lexical decision, different from word pairs with lexically non-contrastive tonal variations (Wu, Chen, Van Heuven, & Schiller, 2014). These findings support that stress and tonal minimal pairs have distinctive representations in lexical access. However, the role of tone in bilingual lexical representation and access needs further investigation.

Another aspect under consideration is the relative frequency of the translation equivalents. Language-dominance effects in lexical access are considered to be mediated by the relative frequencies of lexical representations in the integrated bilingual lexicon (W. J. B. Van Heuven, Dijkstra, & Grainger, 1998). The dominant language is used more frequently in general, and thus activated more easily compared to its translation equivalent. Such language dominance effects are usually taken as granted for normal translation equivalents and reported together with the asymmetrical translation priming effects (Basnight-Brown & Altarriba, 2007) and the asymmetrical cognate facilitation effects (Brenders et al., 2011; Van Hell & Dijkstra, 2002). The frequency-based account for this language dominance effect is built on an important assumption that the translation equivalents have two lexical representations (Altarriba, 1992). Nevertheless, we cannot rule out the possibility that identical translation equivalents are instead represented with one single lexical node. The frequency-based account for the language dominance effect predicts that if a pair of translation equivalents shares a single lexical representation, they cannot be assigned different lexical frequencies and the common lexical node should be activated with the same speed in different language modes, showing no language dominance effect. As previously mentioned, the common representation for the orthographically identical Dutch-English translation equivalents in visual word recognition was supported with empirical evidence (Dijkstra et al., 2010). However, the previous discussion on language-dominance effects rarely considered the identical or near-identical translation equivalents.

In order to study the role of tone in bilingual lexical representation and access of etymologically related translation equivalents, the present study investigated how the cross-linguistic tonal similarity affects the cognate facilitation effect. More specifically, we are looking into whether tonal similarity affects the cognate facilitation effect on segmentally identical translation equivalents and, if it does, whether tonal similarity affects tone-identical (S+T+) and tone-non-identical (S+T−) translation equivalents in the same way.

Different viewpoints with respect to the role of lexical tone in lexical representation provide different answers to these questions. In the discussion of the following three viewpoints as shown in Figure 1, we assume a localist connectionist account for the mental representation of translation equivalents. Under this framework, it is assumed that lexical representations made up by different phonemes are also different within and across languages (Dijkstra et al., 1999). According to this account, there can be a single representation for identical translation equivalents, but at least two representations are needed for non-identical translation equivalents. Dijkstra et al.'s study (2010) supported this account for bilingual visual word recognition. The crucial evidence was that the reaction time sharply increased in nearly identical cognates compared with fully identical cognates, revealing a lateral inhibition effect (Dijkstra et al., 2010). Does this point also apply to the usage of tone in lexical storage by bilinguals who use two closely related tonal dialects rich in segmentally identical translation equivalents? There is evidence indicating that suprasegmental cues on the lexical level are exploited in recognizing spoken words (Cutler & Van Donselaar, 2001; V. J. van Heuven, 1988; Wu et al., 2014). However, does it mean that lexical representations made up by identical segments but different tones across languages/dialects are also different?

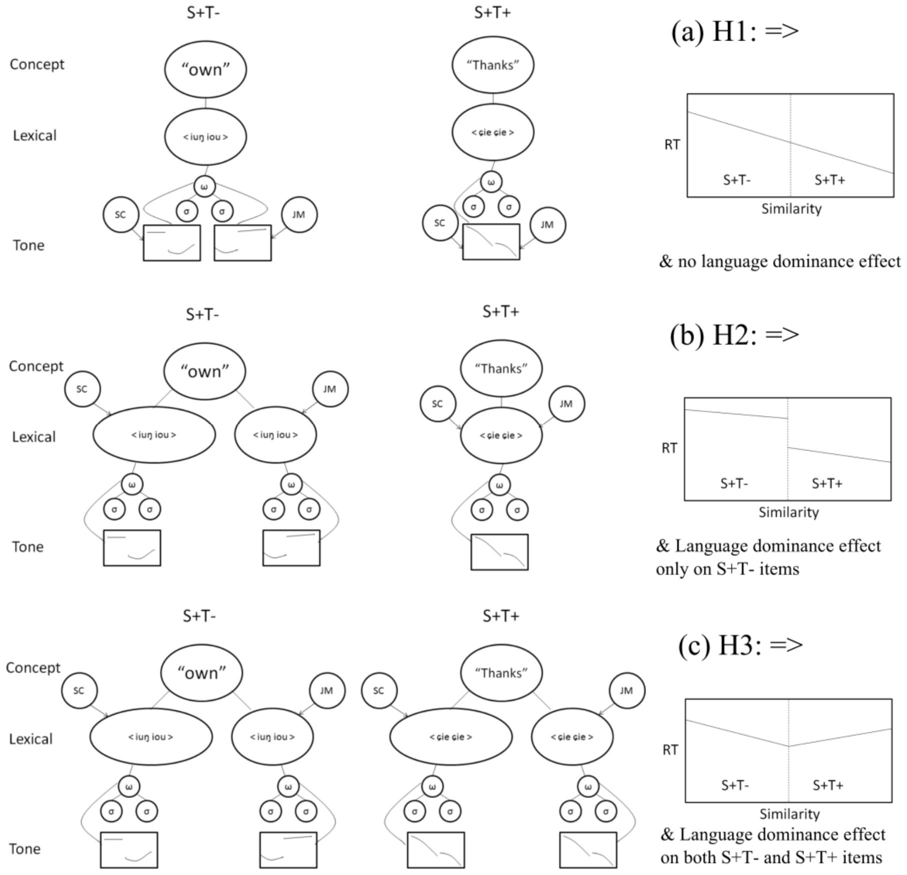


Figure 1. *Three alternative hypotheses on the mental representation of tonally identical (S+T+) and tonally non-identical (S+T−) translation equivalents.*

As mentioned above SC-JM tonal bilinguals speak two closely related tonal dialects, which are teaming with etymologically related translation equivalents. If these tonal bilinguals' lexical representations depend on segmental compositions but tones are assigned later depending on the language mode, S+T+ and S+T− translation equivalents would both be stored as single representations. Then S+T+ translation equivalents could be processed faster than S+T− translation equivalents due to the facilitation from the additional phonological similarity, but the decrease of reaction time with the increase of similarity should be continuous across both types of translation equivalents. Moreover, since language-dominance effects in lexical access are considered to be as mediated by the relative frequencies of lexical representations in the integrated bilingual lexicon (W. J. B. Van Heuven et al., 1998), language-dominance effect should appear on neither type of translation equivalents, as both types are commonly represented.

If SC-JM tonal bilinguals' lexical representations depend on both segmental and tonal compositions but not on the language attribution, S+T+ translation equivalents would be stored as a single representation but S+T− translation equivalents would be stored as two separate representations. Then S+T− translation equivalents would be processed much slower than S+T+ translation equivalents, revealing a similar lateral inhibition effect as found earlier on Dutch-English non-identical cognates (Dijkstra et al., 2010). A general facilitation from tonal similarity could still exist, but the S+T+ translation equivalents should enjoy a large discontinuous processing advantage as found earlier (Dijkstra et al., 2010). Moreover, the language-dominance effect should only appear on the separately represented S+T− translation equivalents but not on the commonly represented S+T+ translation equivalents.

Alternatively, if SC-JM tonal bilinguals' lexical representations not only depend on segmental and tonal compositions but also on the language attribution, not only S+T− but also S+T+ translation equivalents would be stored as two separate representations. In this case, both types of translation equivalents would have lateral inhibition and the difference between S+T− and S+T+ translation equivalents would be very small. Nevertheless, the increase of tonal similarity should yield different effects on the reaction times (RTs) of S+T− and S+T+ items. Since the two representations of S+T+ translation equivalents are nearly identical, we expect to see the lateral inhibition effect increase with the increase of tonal similarity. As for the more different S+T+ translation equivalents, the co-activation facilitation could be more dominant than the lateral inhibition and we would observe increased facilitation with the increase of tonal similarity. Moreover, the language-dominance effect should appear on both S+T− and S+T+ translation equivalents, since both types are represented separately.

4.2 Experiment

4.2.1 Participants

Forty-eight native tonal monolinguals of SC from Beijing (7 male and 41 female, the age ranged from 19 to 30, M = 22.73, SD = 2.95) and 54 native SC-JM tonal bilinguals from Jinan (15 male and 39 female, the age ranged from 19 to 36, M = 22.59, SD = 3.88, 44 SC dominant or balanced, 10 JM dominant) participated in this experiment in exchange for payment. Both groups were right-handed, received their literacy educations in SC, and have learned some English at school. A few participants from each group also have some knowledge of other non-tonal foreign languages, such as French and German.

4.2.2 Design and stimuli

An unbalanced mixed design, since no data was available before the experiment for the phonological similarity between SC-JM translation equivalents, was adopted with Tonal Identity, Tonal Similarity, Word Frequency, Language Mode, and Block as predictors. We first composed a list including 54 pairs of disyllabic SC-JM translation equivalents. Since no measurement of phonological similarity between

SC-JM translation equivalents was available before the experiment, the first author (a trained phonetician with Putonghua Proficiency Test Certificates- Level1B) judged the words from a JM audio corpus (200 high-frequency and 200 low-frequency words by 42 JM speakers collected in our earlier study) for their phonological similarity to their SC counterparts. Then 27 S+T+ and 27 S+T− pairs of translation equivalents were selected (four were later excluded because of segmental differences). Since many JM words were produced with different variants in the corpus, we selected words with dominant-variant probabilities greater than 0.85 and only used the only or highly dominant variants in our experiment. The S+T+ and S+T− candidates were matched with respect to their Chinese word frequency and dominant-variant probability. We also composed a list including 54 pairs of disyllabic non-words using non-existing combinations of Chinese characters. These words and non-words were then produced in both JM and SC by a male native bilingual who is highly proficient in both dialects (also a trained phonetician with Putonghua Proficiency Test Certificates- Level1B). After the main experiment, the translation equivalents were marked again as identical or non-identical according to the similarity rating by both SC monolinguals and SC-JM bilinguals.

The whole SC version of the words and non-words were presented to Beijing tonal monolinguals. The bilinguals were tested in both SC and JM. To eliminate the possibility of within- and between-dialect repetition priming, each bilingual heard only one member of each pair and only heard each stimulus once. The list of pairs was split into two halves (List-A & List-B) with matched number of between-dialect identical candidates, word frequency, dominant-variant probability, style, and tonal categories. Half of the participants heard the SC part of List-A and the JM part of List-B; the other half of the participants heard the SC part of List-B and the JM part of List-A. The SC words and JM words were presented in blocks separated by short breaks. Half of the bilinguals were tested with the SC block first and the other half were tested with the JM block first. Half-lists, language modes, and the test order of language modes were counterbalanced across the bilinguals as shown in Table 1.

Table 1. Counterbalanced design

	JM	SC	Test Order of Language Modes
Bilingual 1	List-A	List-B	JM first
Bilingual 2	List-B	List-A	JM first
Bilingual 3	List-A	List-B	SC first
Bilingual 4	List-B	List-A	SC first

4.2.3 Procedure

Participants were tested individually in a quiet room using the E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). They were told that they would hear a series of sound sequences and they had to decide whether or not each of these sound sequences was a real word. Each item was played binaurally through headphones, with instructions on the screen. A new trial started 1,000 ms after the participant responded to an item, or 1,500 ms after the response time exceeded 5 s. SC and JM words were presented in two blocks separated by a break, in random order. The critical trials of each block were preceded by a practice block including 10 words and 10 non-words.

The language mode was implicitly hinted. At the beginning of each block, the participants heard instructions in the test dialect. Identical translation equivalents were avoided in the practice block. All the trials in one block were in the same dialect, except for the identical translation equivalents, which could be ambiguous regarding the dialect.

After the main experiment, both bilinguals and monolinguals rated all the SC-JM item pairs for cross-linguistic phonological similarity on a five-point scale. Each pair was aurally presented twice to the same participant in two blocks, once with the SC item first and once with the JM item first. The order of SC-first and JM-first presentations was counterbalanced across participants. None of the participants noticed the cross-linguistic tonal similarity of the translation equivalents before the rating phase.

4.3 Analysis and results

4.3.1 Tonal similarity and tonal identity

The similarity rating for each pair was calculated by averaging the ratings across the participants. We compared the average by-item similarity ratings by bilinguals and monolinguals and found a strong by-item correlation, r = .98. By-item paired t-test showed that bilinguals generally rated the pairs as more similar, t (49) = -4.65, p < 0.001. This bias was removed by z-normalizing the mean by-item ratings, t (49) = 0.79, p > 0.05. Since the SC-JM translation equivalents in the present study are segmentally identical but vary in tonal similarity, we treated the rating of phonological similarity by the bilinguals for each item as the Tonal Similarity of the item.

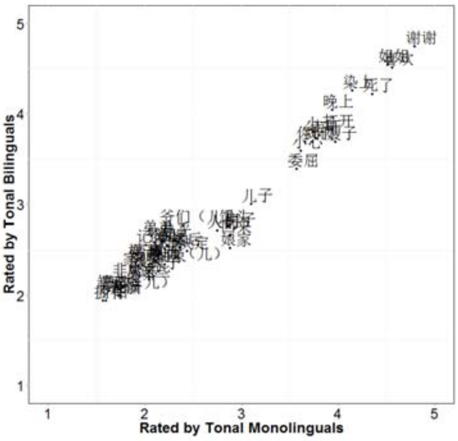


Figure 2. *A scatter plot for the similarity rating (1 = quite dissimilar, 5 = identical) by tonal monolinguals (along the horizontal axis) and tonal bilinguals (along the vertical axis). Each pair of translation equivalents is represented as a point in the plot with their common written form marked.*

The similarity rating showed a smooth dual-peak distribution as shown in Figure 2. We wrote a derivative-based algorithm to detect the virtual bottom of the valley between the two peaks and classified the pairs as tonally identical (S+T+) and tonally non-identical (S+T−) translation equivalents (threshold = 3.30). In the following analysis of RTs, we introduced the factor *Tonal Identity* specifying whether the SC and JM pronunciations are S+T+ or S+T−.

4.3.2 Reaction times

RT analysis was only based on the correct trials. We excluded the 10 JM-dominant bilinguals and one bilingual with a suspicious accent in the analysis of reaction times (RT)[8]. To improve the distribution of the data, RT data were log-transformed (natural log). The RT outliers were excluded for each participant using a distribution based approach (method I) (van der Loo, 2010) on the log transformed RTs, leaving 2143 data points from the monolinguals and 1846 data points from the bilinguals (930 to JM stimuli, 916 to SC stimuli).

In the following analysis, a model was first fitted with all data points, and then in a model criticism we removed data points with standardized residuals exceeding 2.5 standard deviation units from the data set (less than 2.5% of the data) and refitted the model with the trimmed data set. We report the model statistics from the trimmed models.

Since the *Tonal Identity* was a factor derived from the *Tonal Similarity* rating, these two predictors were inherently highly correlated. We first built a set of Linear Mixed Effect (LME) models (Bates, Maechler, Bolker, & Walker, 2013) including only the factorial *Tonal Identity* but not the *Tonal Similarity* in the set of fixed predictors, to investigate the tonal identity effects and its interaction with the other factors. Then we built separate models for S+T+ and S+T− items, including the rating of *Tonal Similarity* and the other fixed predictors, in search of the potential discontinuity of tonal similarity effects. The random terms include by-pair and by-participant intercepts or slopes for the effect of Trial, selected via model comparison based on likelihood ratio tests. The LME models are summarized in the **Appendices** with Satterthwaite approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2013; SAS, 1978).

Language dominance and word frequency effects. As shown in Appendix 1, the first LME model only used the bilinguals' reaction time data and included *Tonal Identity*, *Word Frequency*, *Language Mode*, *Block*, scaled *Trial Order*, and the two-way and three-way interactions of the first four factors as the fixed predictors. The chosen random terms were by-pair random intercepts and by-participant random slopes for the effect of scaled *Trial Order*. The main effects of *Word Frequency*, *Language Mode*, and *Block* and the interaction between *Word*

---

[8] The 10-JM dominant participants seem to show a different pattern in reaction times. However, the variance was too big and we were not able to recruit enough such participants.

Frequency and *Language Mode* were significant. However, the main effect and interactions of *Tonal Identity* were not.

Since the monolinguals listened to all the SC items in one block, two separated LME models were built to compare the bilinguals' RTs in different blocks with the monolinguals' RTs, as shown in Appendices 2 & 3. Both models included *Tonal Identity*, *Word Frequency*, *Language Mode* (in SC by monolinguals, in SC by bilinguals, and in JM by bilinguals), scaled *Trial Order*, and the two-way and three-way interactions of the first three factors as the fixed predictors. The chosen random terms were also by-pair random intercepts and by-participant random slopes for the effect of scaled *Trial Order*. The main effects of *Word Frequency* were significant for both blocks. However, the main effect of *Language Mode* was only significant in the model for the second block but insignificant in the model for the first block. The main effect and interactions of *Tonal Identity* were still insignificant in these two models.

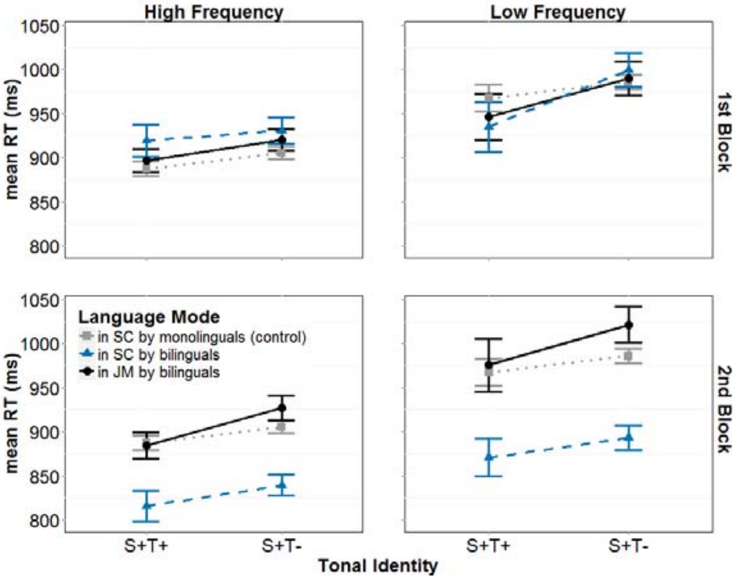


Figure 3. *Mean RT (ms, correct lexical decisions only) for equivalent pairs with and without tonal similarity (S+T+ v.s. S+T−) broken down by three groups of listeners (JM bilinguals, SC bilinguals, SC monolinguals = control). Results are separated for high frequency (left column) and low frequency words (right column) and for first (top row) and second (bottom row) stimulus blocks. Error bars are +/− 1 SE.*

According to the Post-hoc analysis of Differences of Least Squares Means (DLSM) for these models and the descriptive statistics shown in Figure 3, high-frequency words were processed significantly faster than low-frequency words and SC items were processed faster than JM items by the bilinguals in the second block, showing a language dominance effect. The language dominance effect was greater for low-frequency words than for high-frequency words. The bilinguals' RTs in the first block were not different from the monolinguals' RTs but in the second block bilinguals responded significantly faster than the monolinguals to the SC items. Nevertheless, the S+T+ items were not significantly faster than the S+T− items under any condition.

Reversed tonal similarity effects for S+T− and S+T+ items. The following models investigate the discontinuity of tonal similarity effects on tonal bilinguals. Note that the Tonal Similarity ratings used here were scaled and centralized for the S+T+ and S+T− items separately. The general difference of Tonal Similarity between the S+T+ and S+T− items was removed in this analysis.

Two separated LME models were first built for the bilinguals' RTs to the S+T+ and S+T− items, as shown in Appendices 4 &5. Both models included scaled *Tonal Similarity*, *Word Frequency*, *Language Mode*, *Block*, and the two-way, three-way, and four-way interactions of the four factors as the fixed predictors. The chosen random terms for the S+T+ model were by-pair and by-participant random intercepts. The chosen random terms for the S+T− mode were by-pair random intercepts and by-participant random slopes for the effect of scaled *Trial Order*. The main effects of *Language Mode* and *Block* were significant in both models. As for the S+T+ items, the main effects of *Tonal Similarity* and *Word Frequency* were insignificant, but the interaction of *Tonal Similarity*, *Language Mode*, and *Block* was significant. As for the S+T− items, the main effect of *Tonal Similarity* and *Word Frequency* was significant. The two-way interactions of *Tonal Similarity* and *Language Mode*, of *Word Frequency* and *Language Mode*, and the three-way interactions of *Tonal Similarity*, *Language Mode*, and *Block*, of *Word Frequency* and *Language Mode* and *Block* were also significant.

As shown by the model estimates in Figure 4, with S+T+ items, RT increased with the increase of tonal similarity or at least did not decrease with it. *Tonal Similarity* interacted with *Language Mode* differently in the first and second blocks. In the first block, the RT difference introduced by *Language Mode* increased with the increase of *Tonal Similarity*. In the second block, the RT difference introduced by *Language Mode* decreased with the increase of *Tonal Similarity*. Also, SC items were processed faster than their JM counterparts in the second block. S+T− items showed a very different pattern, i.e. RT decreased with the increase of tonal similarity. In the first block, only the *Word Frequency* effect was salient. However, in the second block, not only the effect of *Word Frequency* but also the effect of *Language Mode* was salient. Low-frequency JM words seemed to be the least sensitive to the increase of tonal similarity and the RT difference introduced by *Language Mode* increased with *Tonal Similarity* for low frequency words.

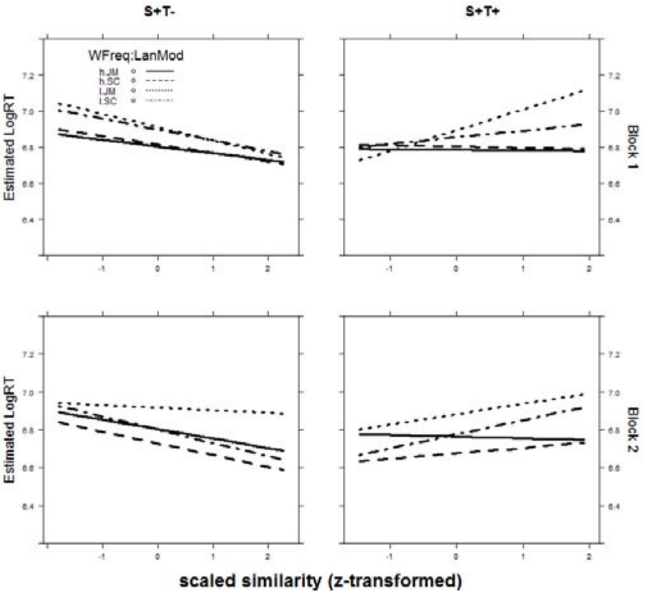


Figure 4. *Mean (estimated) bilingual lexical decision time (natural logarithm, ms) as a function of (z-normalized) judged tonal similarity between SC and JM counterparts presented separately for words with high and low Word Frequency and Language Mode. Linear regression lines are indicated for each combination of Word Frequency and Language Mode. Results are paneled by stimulus block (top row: first block, bottom row: second block) and by absence versus presence of tonal identity (left column: S+T−; right column: S+T+).*

However, similar decrease of RTs with the increase of tonal similarity was also found in the monolinguals' RTs to the S+T− items, F (1) = 17.31, p < 0.001, as shown in Appendices 6 & 7. To investigate the real tonal similarity effect, we compared the effect of *Tonal Similarity* on the bilinguals with the effect obtained for the monolinguals. Since *Block* showed important interactions with the other predictors, two LME models were built for the bilinguals' RTs in the first and second blocks respectively, as shown in Appendices 8 & 9. Both models included scaled *Tonal Similarity*, *Word Frequency*, *Language Mode* (in SC by monolinguals, in SC by bilinguals, and in JM by bilinguals), scaled *Trial Order*, and the two-way and three-way interactions of the first three factors as the fixed predictors. The main effects of *Tonal Similarity* and *Word Frequency* were significant in both the first and the second blocks. However, the interaction of *Tonal Similarity* and *Language Mode* was only significant in the second block.

As shown by the model estimates in Figure 5, although the bilinguals were not different from the monolinguals in the first block of S+T− items, in the second block the bilinguals showed a steeper slope to SC items compared with the slopes of monolinguals, indicating a real facilitation (i.e. shortening the lexical decision time) with the increase of tonal similarity. The bilinguals' slopes to JM items were not different from the slopes by monolinguals in the first block and shallower than the slopes by monolinguals in the second block, indicating an actual interference in RTs with the increase of tonal similarity.

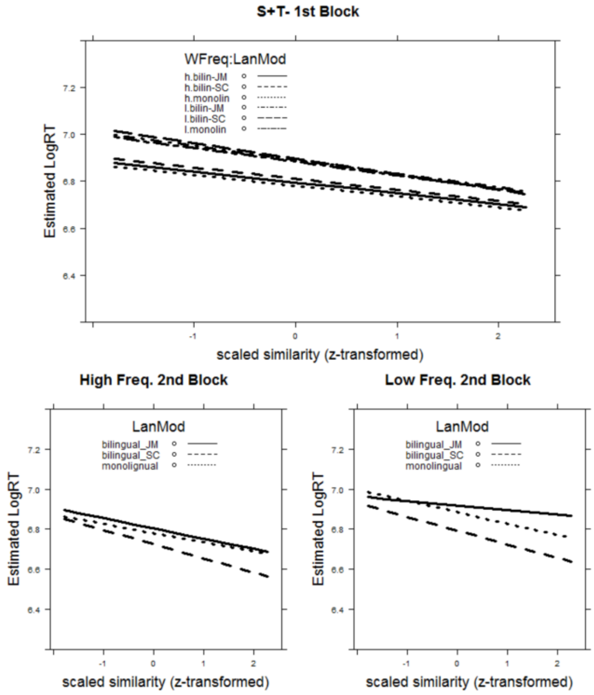


Figure 5. *Mean (estimated) lexical decision time (natural logarithm, ms) obtained for S+T− items as a function of (z-normalized) judged tonal similarity between SC and JM counterparts. Results are paneled by Block (top row: first block, bottom row: second block) and by Word Frequency (left column: high-frequency; right column: low frequency) in the second row. In the upper panel, linear regression lines are indicated for each combination of Word Frequency, Listener Group, and Language Mode (WFreq: h = high, l = low; LanMod: bilin-JM = bilinguals in JM, bilin-SC = bilinguals in SC; monolin = monolinguals in SC). In the lower panel, linear regression lines are indicated for each combination of listener group and Language Mode.*

4.4 Discussion

4.4.1 Main findings

As predicted by the dialect-dependent separate representation hypothesis, the discontinuity of tonal similarity effects and the language-dominance effect were verified on SC-JM tonal bilinguals. First, no significant difference in lexical decision times was found between S+T+ and S+T− translation equivalents. This finding can be understood if we assume that both S+T− and S+T+ translation equivalents are subject to lateral inhibition between two separate lexical nodes. Second, tonal similarity showed different effects on S+T+ and S+T− translation equivalents. Although the main effect was insignificant, the increase of tonal similarity in many conditions increased the reaction times to the S+T+ items, revealing a prevalence of lateral inhibition effects. However, the increase of tonal similarity instead reduced the reaction times to the S+T− items, revealing a prevalence of facilitatory coactivation. Third, and more importantly, a language dominance effect was found on the SC words against their JM counterparts in the second block. Such a language dominance effect should only be found for translation equivalents, if these are stored separately. All these findings are consistent with the predictions from the dialect-dependent separate representation hypothesis.

Tonal similarity also showed different complex interactions with *Language Mode* and *Block* on S+T+ and S+T− translation equivalents, some of which were not predicted by any of the candidate hypotheses. Here we first consider the interactions in the second block, where a general language dominance effect was found (see the lower plots in Figure 3). The increase of tonal similarity reduced the RT difference between the SC and JM translation equivalents if they were tone identical (S+T+) but increased such RT difference if the translation equivalents were tonally non-identical (S+T−) (see the interaction of *Language Mode* and *Tonal Similarity* in the lower panel of Figure 4 and the lower panel of Figure 5.) The RT difference between the SC and JM translation equivalents reveals a language dominance effect. Thus, language dominance effects are modulated by tonal similarity in different ways with the S+T+ and S+T− translation equivalents. These findings again suggest that tonal similarity influences the bilingual lexical access of S+T+ and S+T− translation equivalents in different ways and are in line with the claim that both types of translation equivalents should be stored as two separate representations. Looking at such interactions from another perspective, the interaction with *Language Mode* plays an important role in modulating the way *Tonal Similarity* takes effect. With S+T+ translation equivalents, although the processing of the other types of items was negatively affected by the increase of tonal similarity, the processing of high-frequency JM items was not sensitive to tonal similarity at all (see the lower right panel in Figure 4). Also with S+T− translation equivalents, further comparison with monolinguals showed that tonal similarity actually only facilitated the bilinguals' responses to the SC items but interfered with the bilinguals' responses to the JM items (see the lower right panel in Figure 5). Taken together, with S+T+ translation equivalents, tonal similarity reduces the language dominance effect and interferes with the lexical access of SC words and low-frequency JM words. With S+T− translation equivalents, tonal similarity strengthens the language dominance effect, facilitates the lexical access of SC words, but interferes with the lexical access of low-frequency JM words; and the lexical access of high-frequency JM words is not sensitive to tonal similarity. Note that these findings only apply to the second block. In the first block, the S+T+ translation equivalents showed a reversed interaction of *Language Mode* and *Tonal Similarity* (strengthens the language dominance effect) on the low frequency words and showed no effect or interaction on the high frequency words; and the S+T− translation equivalents were only sensitive to *Word Frequency*. The theoretical interpretations for these interactions are provided in the following section.

4.4.2 Theoretical implications

Just like lexical stress (Levelt et al., 1999; Schiller & Costa, 2006), lexical tone functions like abstract lexical frame. Just like minimal stress pairs (Cutler & Van Donselaar, 2001), tonal minimal pairs have distinctive lexical representations (Wu et al., 2014). Lexical tone not only distinguishes lexical representations with different meanings within one language (Wu et al., 2014) but also distinguishes the lexical representations of translation equivalents in the bilingual lexicons, which is supported by the current finding that language dominance affects the reaction times of translation equivalents which are only different in tone. Moreover, it seems that language mode itself is sufficient for distinguishing bilingual lexical representations in the current case of bilingualism because even when the translation equivalents were not only segmentally but also tonally identical, the language dominance effect persisted.

The language dominance effects we found support separated-representations for all the orthographically identical translation equivalents. Since both S+T+ and S+T− stimuli in the current study are comparable to the orthographically identical Dutch-English cognates used in (Dijkstra et al., 2010), our claim seems inconsistent with the earlier claim that the orthographically identical cognates share one common lexical representation (Dijkstra et al., 2010). However, this inconsistency may be attributed to two causes. First, the previous study used visual stimuli and the current study used auditory stimuli. The inconsistent findings invite us to consider the possibility that the underlying lexical representations accessed via the auditory route may differ from those accessed via the visual route. Second, Dutch and English are two different languages but SC and JM are two closely related tonal dialects. As shown in a previous study (Lemhöfer et al., 2008), the different cognate status of translation equivalents in different language combinations affect L2 word identification. It is reasonable to believe that a bilingual lexicon dominated by etymologically related translation equivalents would function differently from a bilingual lexicon where etymologically related translation equivalents (cognates) only exist sporadically. The bilingualism involving closely related dialects needs further investigation.

The language dominance effect in the current study is related but not directly comparable to the asymmetrical cognate facilitation effects found in earlier studies. L2 words with L1 cognates were more likely to show cognate facilitation effects than vice versa (Brenders et al., 2011) and L1 words with L2 cognates were more likely to show cognate facilitation effects than L1 words with L3 cognates (Van Hell & Dijkstra, 2002). Both studies used language-specific words as controls. However, in the current study too few unrelated control words were available and hence were not included. Instead, the monolinguals' reaction times to the same SC words were used as monolingual controls. The difference in the type of control makes the current results not directly comparable to the earlier findings. The tonal bilinguals' reactions to the JM words were not faster than the monolinguals' reaction times to the SC counterparts. However, the bilinguals' reactions to the SC words were even faster than the monolinguals'. Thus, if what we found counts as cognate facilitations, the cognate facilitation is stronger on the dominant dialect SC, which is different from the earlier findings where the cognate facilitation was stronger on L2. Such difference can again be attributed to the different structure of bilingual mental lexicon as previously discussed. Nevertheless, the SC-JM bilinguals are simultaneous bilinguals and both dialects are their L1. The difference of sequential and simultaneous bilingualism can be another resource of the different direction of asymmetry.

The language dominance effect found on these tonal bilinguals even made the same SC words accessed faster by the bilinguals than by the monolinguals, yielding an unusual bilingual lexical advantage. Most previous studies showed that bilinguals are slower in lexical access compared with monolinguals (Bialystok, 2009; Martin et al., 2012; Ransdell & Fischler, 1987). The bilingual lexical disadvantage was explained by the way that bilinguals have a denser lexical neighborhood and hence suffer from more lateral inhibitions than monolinguals (Ransdell & Fischler, 1987). This mechanism is also verified by computer simulation: just like adding words from the original language, adding words from a new language increases reaction times (Dijkstra, 2003). More importantly, even with the same cognates, the bilinguals were still mostly found to be slower than the monolinguals in cognate production (Costa et al., 2000) and in visual word recognition (Dijkstra et al., 1999; Lemhöfer, Dijkstra, & Michel, 2004; Lemhöfer et al., 2008; Mulder, Dijkstra, Schreuder, & Baayen, 2014). This is not surprising because in the cases of bilingualism studied previously, cognates are usually orthographically and phonologically non-identical and hence the neighborhoods are also denser than those of the monolinguals. However, as discussed in the previous paragraphs, the SC-JM bilingual lexicon is dominated by orthographically identical and phonologically related translation equivalents. Moreover, the language dominance effect on the translation equivalents support that the SC and JM translation equivalents are stored as separated lexical nodes. Thus, in a similarly dense neighborhood, the two lexical nodes for the SC-JM translation equivalents may interact in a different way compared with the separated lexical nodes stored in the Dutch-English bilingual lexicon. The segmentally identical translation equivalents in the JM-SC bilingual lexicon may benefit more from the coactivation and suffer less from the lateral inhibition. Furthermore, the facilitatory coactivation may be in dominance in the SC-JM bilingual lexicon (with tonal similarity providing minute adjustments on this basis) and provide the bilinguals some advantage in lexical access compared with the monolinguals.

Tonal similarity showed discontinuous effects on the S+T+ and S+T− translation equivalents. The discontinuous effects of tonal similarity are in line with the earlier findings. Dijkstra et al. (2010) also found discontinuous effects of phonological similarity. The increase of phonological similarity facilitated the lexical decision of orthographically identical Dutch-English cognates but not on the orthographically non-identical cognates (Dijkstra et al., 2010). The discontinuity was attributed to the lateral inhibition effects introduced by the additional lexical representation of the orthographically near-identical cognates (Dijkstra et al., 2010). Lateral inhibition increases with the increase of similarity between the representations concerned. Indeed, in the current study, with the S+T+ translation equivalents, we found reaction times increased with the increase of tonal similarity on the SC words, revealing a dominant lateral inhibition effect. However, such effects were scarce on the comparable S+T− translation equivalents, where facilitatory coactivation was dominant instead.

Moreover, the discontinuous interactions of tonal similarity with language modes suggest that lexical nodes from the dominant dialect may be more sensitive to both the facilitatory coactivation and the lateral inhibition. The RT difference between a pair of translation equivalents reveals language dominance effect (W. J. B. Van Heuven et al., 1998). The SC-dominant tonal bilinguals responded faster to the SC words than their JM counterparts because they use the SC versions more often. However, the increase of tonal similarity reduced the language dominance effect on S+T+ translation equivalents but enhanced the language dominance effect on S+T− translation equivalents. One possible interpretation is that the translation equivalents with smaller language dominance effect have more balanced relative frequency of the SC and JM forms. However, there is no reason to assume a relation between the between-dialect tonal similarity and the relative frequency of word usage. Another interpretation may be more reasonable. Lexical nodes from the dominant dialect SC are more sensitive to both the facilitatory coactivation and the lateral inhibition than their counterparts from the less dominant dialect JM. As a result, with the S+T+ translation equivalents, where the reaction times increase with the increase of tonal similarity, the reaction times to the SC words increased faster and approached the reaction times to their JM counterparts; and with the S+T− translation equivalents, where the reaction times decrease with the increase of tonal similarity, the reaction times to the SC words decreased faster and deviated even more from their JM counterparts.

Most of the above-mentioned findings were only found in the second block. In the first block, except for some general word frequency effects, no language dominance effect and not many tonal similarity effects were found. Such block effects need to be put in context of bilingual cognitive control.

The differences between the first and second block seem to be more related to the bilinguals' general control of attention. Previous studies have shown that bilingual lexical access may be more or less language-selective depending on the language mode (language-specific vs. general) (Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000) and the construction of list (mixed vs. monolingual) (Brenders et al., 2011; Caramazza & Brones, 1979; Dijkstra, De Bruijn et al., 2000; Dijkstra, Timmermans, & Schriefers, 2000). In the current study, the language mode was implicitly hinted and the participants first came across a monolingual list and then switched to the other dialect. It could be that the bilinguals' attention was first directed to the first-appeared dialect alone and tuned into a more selective mode of lexical access until the second block started. When the second block started, the bilinguals noticed that the dialect changed and they were in a bilingual situation, they became less selective and that is why most of the bilingual effects emerged in the second block.

Two alternative explanations can be ruled out. First, the block asymmetry is not due to asymmetrical translation priming. Translation equivalents can prime each other (Duyck & Warlop, 2009) but forward translation priming (dominant language to non-dominant language or L1 to L2) is more robust than backward translation priming (non-dominant language to dominant language or L2 to L1) (Alvarez, Holcomb, & Grainger, 2003; Finkbeiner, Forster, Nicol, & Nakamura, 2004; Gollan, Forster, & Frost, 1997; Midgley, Holcomb, & Grainger, 2009). Such asymmetrical translation priming is not applicable in the current study because the two members of the same pair of translation equivalents were never presented to the same participant in the current study.

Second, since the second block is always in a different dialect compared to the first block, the current block-dependent asymmetry may relate to the asymmetrical switching costs. Switching from the non-dominant language to the dominant language causes greater switching cost than vice versa in speech production (Costa & Santesteban, 2004; Meuter & Allport, 1999). Similar asymmetrical global switching costs between different language blocks were also found when bilinguals name the same set of pictures first in L1 and then in L2 (Guo, Liu, Misra, & Kroll, 2011). However, when cognates were embedded in monolingual sentences, neither local switching cost was found when switching the language of the sentence, nor different magnitudes of cognate facilitation were found between blocks when the language of the sentences was blocked (Gullifer, Kroll, & Dussias, 2013). The blocking of the current study is similar to the case of global switching (Guo et al., 2011), except that it taps into recognition instead of production and bilinguals never heard the two members of the same pair of translation equivalent repeated. However, our finding is still different. Instead of observing greater switching costs in the block of the dominant dialect (Guo et al., 2011) or null effect of global switching (Gullifer et al., 2013), words from the dominant dialects SC were not only processed faster than their JM counterparts in the second block but also faster than the same SC words processed by the monolinguals controls. Thus what we found is a language dominance effect and a bilingual lexical advantage but not a classical asymmetrical switching cost.

In sum, the new findings of language dominance effects and bilingual lexical advantage by the SC-JM tonal bilinguals remind us to pay more attention to the structure of the bilingual lexicon. A bilingual lexicon filled with etymologically related translation equivalents may be organized and function differently from a bilingual lexicon dominated with etymologically irrelevant translation equivalents. The new findings of discontinuous tonal similarity effects and its interaction with the language dominance effect provide us further insights into the role of lexical tones in the bilingual lexical representation and lexical access. The strengths of facilitatory coactivation and lateral inhibition may be not only related to the similarity of the translation equivalents but also dynamically modulated by the language mode and the bilinguals' attention. These findings together with the block effect also suggest that bilingual lexical access may be more or less language-selective depending on the bilinguals' language environment.

References

Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. *Advances in Psychology, 83*, 157-174. doi: 10.1016/S0166-4115(08)61493-4

Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meaning in two languages: An event-related brain potential study of beginning bilinguals. *Brain and Language, 87*(2), 290-304. doi: 10.1016/S0093-934X(03)00108-1

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35*(5), 953-965.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11. doi: 10.1017/S1366728908003477

Brenders, P., van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of experimental child psychology, 109*(4), 383-396. doi: 10.1016/j.jecp.2011.03.012

Bultena, S., Dijkstra, T., & van Hell, J. G. (2012). Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes*(ahead-of-print), 1-28. doi: 10.1080/01690965.2012.718353

Bultena, S., Dijkstra, T., & van Hell, J. G. (2013). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *The Quarterly Journal of Experimental Psychology*(just-accepted), 1-65. doi: 10.1080/17470218.2013.853090

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society, 13*(4), 212-214. doi: 10.3758/BF03335062

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751-781. doi: 10.1006/jmla.2001.2825

Chen, Y., Shen, R., & Schiller, N. O. (2011). *Representation of allophonic tone sandhi variants.* Paper presented at the Proceedings of Psycholinguistics Representation of Tone. Satellite Workshop to ICPhS, Hongkong.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283. doi: 10.1037/0278-7393.26.5.1283

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*(4), 491-511. doi: 10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language, 94*(1), 94-103. doi: 10.1016/j.bandl.2004.12.002

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology, 38*(3), 367-393. doi: 10.1080/14640748608401604

Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and speech, 44*(2), 171-195. doi: 10.1177/00238309010440020301

Dijkstra, T. (2003). Lexical processing in bilinguals and multilinguals: The word selection problem *The multilingual lexicon* (pp. 11-26): Springer.

Dijkstra, T., De Bruijn, E., Schriefers, H., & Ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism Language and Cognition, 3*(1), 69-78. doi: 10.1017/S1366728900000146

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301. doi: 10.1016/j.jml.2009.12.003

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language, 42*(4), 445-464. doi: 10.1006/jmla.1999.2697

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 663. doi: 10.1037/0278-7393.33.4.663

Duyck, W., & Warlop, N. (2009). Translation priming between the native language and a second language: New evidence from Dutch-French bilinguals. *Experimental Psychology, 56*(3), 173. doi: 10.1027/1618-3169.56.3.173

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51*(1), 1-22. doi: 10.1016/j.jml.2004.01.004

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1122. doi: 10.1037/0278-7393.23.5.1122

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: lexical access in sentence context. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00278

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *Neuroimage*. doi: 10.1016/j.neuroimage.2011.03.049

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106*(1), 501-511. doi: 10.1016/j.cognition.2007.02.001

Jongenburger, W. (1996). *The role of lexical stress during spoken-word processing (Ph.D. dissertation, University of Leiden.)*: Holland Academic Graphics.

Kootstra, G. J., van Hell, J. G., & Dijkstra, T. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency.

Bilingualism: Language and Cognition, 1(1), 1-1. doi: 10.1017/S136672891100068X

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32*(4), 533-550. doi: 10.3758/BF03195845

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*(5), 585-611. doi: 10.1080/01690960444000007

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: a megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 12. doi: 10.1037/0278-7393.34.1.12

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75. doi: 10.1017/S0140525X99001776

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., & Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain and Language, 120*(1), 61-65. doi: 10.1016/j.bandl.2011.10.003

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126. doi: 10.1207/s15516709cog0000_79

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language, 40*(1), 25-40. doi: 10.1006/jmla.1998.2602

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology, 46*(3), 551-565. doi: 10.1111/j.1469-8986.2009.00784.x

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological Abstraction in Processing Lexical - Tone Variation: Evidence From a Learning Paradigm. *Cognitive Science, 35*(1), 184-197. doi: 10.1111/j.1551-6709.2010.01140.x

Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, H. R. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language, 72*, 59-84. doi: 10.1016/j.jml.2013.12.004

Nixon, J. S., Chen, Y., & Schiller, N. O. (2014). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. *Language, Cognition and Neuroscience*(ahead-of-print), 1-15. doi: 10.1080/23273798.2014.942326

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26*(4), 392-405. doi: 10.1016/0749-596X(87)90098-2

SAS. (1978). Tests of Hypotheses in Fixed-Effects Linear Models *SAS Technical Report R-101* Cary, NC, USA: SAS Institute Inc.

Schiller, N. O., & Costa, A. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *The Mental Lexicon, 1*(2), 231-250. doi: 10.1075/ml.1.2.04sch

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

van der Loo, M. P. J. (2010). Distribution based outlier detection for univariate data,. *Discussion paper 10003, Statistics Netherlands, The Hague*.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic bulletin & review, 9*(4), 780-789. doi: 10.3758/BF03196335

van Heuven, V. J. (1988). *Effects of stress and accent on the human recognition of word fragments in spoken context: gating and shadowing.* Paper presented at the Proceedings of the 7th FASE/Speech-88 Symposium, The Institute of Acoustics, Edinburgh.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language, 39*(3), 458-483. doi: 10.1006/jmla.1998.2584

Wu, J., Chen, Y., Van Heuven, V. J., & Schiller, N. O. (2014). Tonal variability in lexical access. *Language, Cognition and Neuroscience*. doi: 10.1080/23273798.2014.915977

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes, 24*(7-8), 1082-1096. doi: 10.1080/01690960802042133

Zhao, H., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: New evidence from picture naming in Chinese. *Acta psychologica, 139*(2), 272-280. doi: 10.1016/j.actpsy.2011.12.001

Appendices

Appendix 1. The model investigating the effects on the bilinguals

Fixed effects	Df	F	p
Tonal Identity	45.43	0.92	> 0.05 (ns)
Word Frequency	45.45	6.96	< 0.05
Language Mode (JM/ SC)	1683.73	53.77	< 0.001
Block	1685.13	46.03	< 0.001
Scaled Trial Order	38.2	0.93	> 0.05 (ns)
Tonal Identity × Word Frequency	45.43	0.04	> 0.05 (ns)
Tonal Identity × Language Mode	1683.58	0.26	> 0.05 (ns)
Word Frequency × Language Mode	1697.71	6.60	< 0.05
Tonal Identity × Block	1679.9	0.65	> 0.05 (ns)
Word Frequency × Block	1695.77	2.20	> 0.05 (ns)
Language Mode × Block	41.77	2.10	> 0.05 (ns)
Tonal Identity × Word Frequency × Language Mode	1692.27	0.038	> 0.05 (ns)
Tonal Identity × Word Frequency × Block	1691.18	0.863	> 0.05 (ns)
Tonal Identity × Language Mode × Block	1679.99	0.096	> 0.05 (ns)
Word Frequency × Language Mode × Block	1687.32	0.503	> 0.05 (ns)
Tonal Identity × Word Frequency × Language Mode × Block	1682.15	0.015	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		28.93	<0.001
1 | pair		679.12	<0.001

Appendix 2. The model investigating the effects in Block 1

Fixed effects	Df	F	p
Tonal Identity	47.26	0.61	> 0.05 (ns)
Word Frequency	47.25	9.64	< 0.01
Language Mode (in SC by monolingual/ in SC by bilingual/ in JM by bilingual)	94.21	0.01	> 0.05 (ns)
Scaled Trial Order	71.38	9.05	< 0.01
Tonal Identity × Word Frequency	47.22	0.05	> 0.05 (ns)
Tonal Identity × Language Mode	2802.56	0.41	> 0.05 (ns)
Word Frequency × Language Mode	2825.11	2.56	> 0.05 (ns)
Tonal Identity × Word Frequency × Language Mode	2818.32	0.32	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		69.41	<0.001
1 | pair		993.96	<0.001

Appendix 3. The model investigating the effects in Block 2

Fixed effects	Df	F	p
Tonal Identity	47.23	1.32	> 0.05 (ns)
Word Frequency	47.21	11.06	< 0.01
Language Mode (in SC by monolingual/ in SC by bilingual/ in JM by bilingual)	95.84	5.63	< 0.01
Scaled Trial Order	63.75	0.81	> 0.05 (ns)
Tonal Identity × Word Frequency	47.16	0.01	> 0.05 (ns)
Tonal Identity × Language Mode	2812.86	2.3	> 0.05 (ns)
Word Frequency × Language Mode	2836.11	2.91	> 0.05 (ns)
Tonal Identity × Word Frequency × Language Mode	2818.39	0.12	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		77.1	<0.001
1 | pair		966.22	<0.001

Appendix 4. The model investigating the effect of Tonal Similarity on S + T + items by the bilinguals

Fixed effects	Df	F	p
scaled Tonal Similarity	10.12	2.16	> 0.05 (ns)
Word Frequency	10.2	4.00	> 0.05 (ns)
Language Mode	487.83	19.31	< 0.001
Block	487.63	23.58	< 0.001
scaled Tonal Similarity × Word Frequency	10.12	1.86	> 0.05 (ns)
scaled Tonal Similarity × Language Mode	505.89	0.23	> 0.05 (ns)
Word Frequency × Language Mode	486.49	2.28	> 0.05 (ns)
scaled Tonal Similarity × Block	506.44	0.02	> 0.05 (ns)
Word Frequency × Block	485.81	1.46	> 0.05 (ns)
Language Mode × Block	44.62	1.87	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Language Mode	493.41	3.26	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Block	493.88	1.01	> 0.05 (ns)
scaled Tonal Similarity × Language Mode × Block	482.3	7.89	<0.01
Word Frequency × Language Mode × Block	483.4	0.65	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Language Mode × Block	482.4	1.3	> 0.05 (ns)
Random effects		χ^2	
1 |participant		178.15	<0.001
1 | pair		109.9	<0.001

Appendix 5. The model investigating the effect of Tonal Similarity on S + T− items by the bilinguals

Fixed effects	Df	F	p
scaled Tonal Similarity	31.93	10.39	< 0.01
Word Frequency	31.53	8.46	< 0.01
Language Mode	1156.74	45	< 0.001
Block	1155.82	37.95	< 0.001
scaled Tonal Similarity × Word Frequency	31.91	0.03	> 0.05 (ns)
scaled Tonal Similarity × Language Mode	1165.66	4.56	< 0.05
Word Frequency × Language Mode	1164.52	5.15	< 0.05
scaled Tonal Similarity × Block	1152.41	0.54	> 0.05 (ns)
Word Frequency × Block	1160.53	0.02	> 0.05 (ns)
Language Mode × Block	39.87	2.61	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Language Mode	1187.86	0.43	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Block	1180.69	5.8	< 0.05
scaled Tonal Similarity × Language Mode × Block	1161.96	6.04	< 0.05
Word Frequency × Language Mode × Block	1154.31	0.39	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Language Mode × Block	1162.22	5.35	< 0.05
Random effects		χ^2	
1 + Scaled Trial Order |participant		15.87	<0.001
1 | pair		397.03	<0.001

Appendix 6. The model investigating the effect of Tonal Similarity on S + T + items by the monolingual controls

Fixed effects	Df	F	p
scaled Tonal Similarity	10.05	0.82	> 0.05 (ns)
Word Frequency	10.1	6.76	< 0.05
Scaled Trial Order	44.81	7.32	< 0.01
scaled Tonal Similarity × Word Frequency	10.04	0.7	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		17.92	<0.001
1 | pair		127.33	<0.001

Appendix 7. The model investigating the effect of Tonal Similarity on S + T− items by the monolingual controls

Fixed effects	Df	F	p
scaled Tonal Similarity	31.38	17.31	< 0.001
Word Frequency	31.16	16.84	< 0.001
scaled Tonal Similarity × Word Frequency	31.4	0.25	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		28.08	< 0.001
1 | pair		287.58	< 0.001

Appendix 8. The model investigating the effect of Tonal Similarity on S+T− items in the first Block

Fixed effects	Df	F	p
scaled Tonal Similarity	33.93	15.84	< 0.001
Word Frequency	33.39	12.06	< 0.01
Language Mode	91.46	0.04	> 0.05 (ns)
Scaled Trial Order	67.98	4.62	<0.05
scaled Tonal Similarity × Word Frequency	34.78	0.27	> 0.05 (ns)
scaled Tonal Similarity × Language Mode	1952.56	0.23	> 0.05 (ns)
Word Frequency × Language Mode	1939.34	0.65	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency × Language Mode	2008.74	0.05	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		24.48	< 0.001
1 | pair		505.10	< 0.001

Appendix 9. The model investigating the effect of Tonal Similarity on S+T− items in the second Block

Fixed effects	Df	F	p
scaled Tonal Similarity	33.54	15.06	< 0.001
Word Frequency	33.17	12.12	< 0.01
Language Mode	93.23	4.49	< 0.05
Scaled Trial Order	61.84	0.22	> 0.05 (ns)
scaled Tonal Similarity × Word Frequency	34.3	0.03	> 0.05 (ns)
scaled Tonal Similarity × Language Mode	1982.16	4.76	< 0.01
Word Frequency × Language Mode	1986.52	4.05	< 0.05
scaled Tonal Similarity × Word Frequency × Language Mode	2017.78	2.08	> 0.05 (ns)
Random effects		χ^2	
1 + Scaled Trial Order |participant		28.65	< 0.001
1 | pair		469.26	< 0.001

5 Tonal Variability in Lexical Access[9]

Abstract

How do different types of tonal variability contribute to lexical access? We addressed this question by investigating a type of variability in Jinan tonal patterns, which is lexically non-contrastive but potentially contrastive in other words. This variability was tested against three levels of variability, viz. ‘acoustic identity’, ‘within-category variation’, and ‘lexically-contrastive variation’, in an auditory lexical decision task. The tonal pattern variation induced a similar but smaller facilitation effect compared with the acoustic identity and the within-category variation. In contrast, an inhibition effect was induced by the lexically-contrastive condition. Additionally, we tested the participants’ tonal awareness. The effect of tonal awareness was smaller on the targets than on the primes. We conclude that, in lexical access, tonal patterns may have representative status but can converge in a lexically-specific way, and that the contribution of tonal awareness is reduced when the form is repeated.

5.1 Introduction

Listeners need to handle different types of variability in speech processing in order to access the correct lexical item, selectively relying on or ignoring the differences between acoustic signals. Most previous work on the processing of variability in lexical access has studied segmental alternations (e.g. Mitterer & Blomert, 2003), or the variability within a specific phoneme category (e.g. Andruski, Blumstein, & Burton, 1994). While much has been learned from studies on segmental variability, it is important to note that many languages make use of suprasegmental properties to signal lexical differences. For instance, pitch works as an important cue in the perception of lexical stress in English (e.g. Fry, 1958). Pitch is also crucial for the distinction of lexical tones in tonal languages. Sixty to 70% of the world’s languages systematically use pitch variation to distinguish lexical meanings (Yip, 2002).

However, the limited number of studies on the processing of tone suggests that tones are processed differently from segments in lexical access. In implicit priming, for instance, in contrast to the traditional facilitation effect observed due to segmental primes, the overlap of Mandarin tones alone (J. Y. Chen, Chen, & Dell, 2002) and the surface tonal overlap accompanied with segmental sharing (Y. Chen, Shen, & Schiller, 2011) produced no facilitatory priming effect in a speech production task. During phonological encoding, reaction time in a phoneme monitoring task (Ye & Connine, 1999) and reaction time and onset latency of the N200 component in a go/nogo task (Zhang & Damian, 2009) showed that segmental information became available prior to tonal information in Mandarin. This difference in priority might also hold for the phonological decoding of auditory

---

[9] This chapter is based on Wu, J., Chen, Y., Van Heuven, V. J., & Schiller, N. O. (2014). Tonal variability in lexical access. *29*(10), 1317-1324. *Language, Cognition and Neuroscience*. doi: 10.1080/23273798.2014.915977

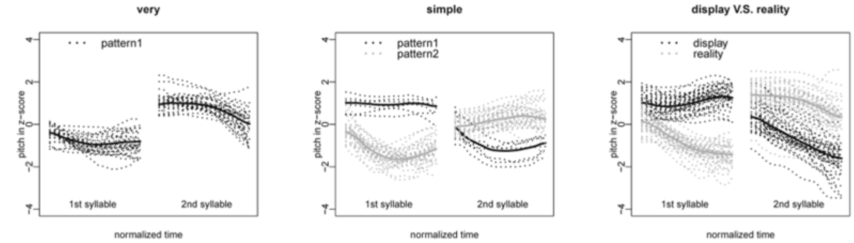
lexical access (Cutler & Chen, 1997). Although lexical adaptation seems to work similarly in tones and consonants (McQueen, Cutler, & Norris, 2006; Mitterer, Chen, & Zhou, 2011), more research is needed to understand the role of tonal variability in lexical access.

We will address the issue with data on tonal variability in Jinan Mandarin, a northern Mandarin dialect of Chinese. Jinan Mandarin has four monosyllabic tones traditionally described as Low-rising, High-falling, High-level, and Low-falling (Qian, 1995). In Jinan, the following type of variability has received relatively little attention in the literature. The same compound word can be accessed through two forms, with the same segmental structure but with distinctive tonal patterns. For instance, the same segmental structure /tɕiɛn tan/ with two different tonal patterns, either HL or LH, provides access to the same word 'simple'. This may sound similar to the phonological variants, such as the flapped /t/ and released /t/, tested earlier (Connine, 2004; Mitterer & Ernestus, 2006). However, unlike the two allophonic variants of /t/, the two tonal patterns of the Jinan word 'simple' are potentially contrastive for certain other segment strings. For instance, /ʂou tɕi/, with the HL tonal contour, provides access to the lexical item 'cell-phone' However, with the LH tonal pattern, it provides access to 'collect'. Thus, if the lack of contrast of the two different tonal patterns from specific words (such as /tɕiɛn tan/) generalizes to the whole vocabulary (including words like /ʂou tɕi/), this could result in confusion regarding many tonal minimal pairs. It is worth noting that the two tonal patterns exist within the same speaker. This is different from the dialectal differences such as American English *tomAYto* versus British English *tomAHto*.

Jinan also has two other types of tonal variability, which are common across the world's languages. One is non-contrastive within-category variation, where two renditions of the same word 'very' can be realized with slight differences in the shape of the pitch contour, although their tonal pattern remains the same. The other is lexically contrastive variation which involves two members of a tonal minimal pair, such as /ɕiɛn ʂʅ/ (HL) 'display'versus /ɕiɛn ʂʅ/ (LH) 'reality'.

Figure 1 illustrates these three types of tonal variability. The tonal pattern variation is similar to contrastive variation and different from within-category variation in that it involves different tonal patterns. However, the two tonal patterns in tonal pattern variation provide access to the same word. This makes it similar to within-category variation but different from contrastive variation.

How this tonal pattern variation is processed in lexical access remains an open question. Are the two forms treated like two different words similar in sound; or are they treated like two renditions of the same word different only in acoustic details? To shed light on these questions, one useful approach is to pit the tonal pattern variation against the other types of tonal variability in the process of lexical access. In the present study, we compared the possible priming effect of (1) acoustic identity, (2) non-contrastive within-category variation, (3) lexically non-contrastive tonal pattern variation, and (4) lexically contrastive tonal variability in an auditory priming paradigm employing the lexical decision task. If the difference between any two types of tonal variability leads to a difference in size of the priming effect, be it facilitation or interference, the two types are likely to be distinguished in their level of perceptual process before their final lexical access.



(a) Within-category variation (single-pattern, non-contrastive) (b) Tonal pattern variation (double-pattern, non-contrastive) (c) Lexically-contrastive variation (double-pattern, contrastive)

Figure 1. *Illustrations of the three types of variability. Each plot was made with recordings produced by 42 Jinan speakers (1 or 2 outliers excluded). Pitch values were z-transformed semitones (the mean and SD were calculated with about 600 recordings for each speaker).*

Based on the reports that tone-mismatch primes lead to inhibition in associative priming (Zhou, Q, Shu, Gaskell, & Marslen-Wilson, 2004), the contrastive variation condition (Condition 4) in the current study is also expected to yield an inhibition effect. The other three conditions are lexically non-contrastive (including the acoustic identity). If non-contrastive tonal variants converge into one representation at some stage of lexical access, both within-category variation and tonal variability should have a priming effect similar to that of the acoustic identity condition, given that both the target and the prime are linked to the same word in all of these conditions. However, if tonal patterns play a role at this stage of lexical access, the amount of priming of tonal pattern variation should differ from within-category variation, given their different quantity and quality of variability. Specifically, the priming effect of tonal pattern variation should be mitigated given that it involves two different tonal patterns.

In addition, general cognitive skills might also affect task performance and individual variation. For instance, tonal awareness, as a subset of phonological awareness, reveals listeners' aptitudes for discriminating and identifying tones (X. Chen et al., 2004; Shu, Peng, & McBride - Chang, 2008). In the current study, this was assessed as a between-participant factor through a tonal oddity task. We introduced this test for two reasons. Firstly, the phonological awareness level may help to explain possible individual variation in the priming/inhibition effect. Secondly, we are interested in the potential differences and interactions between the two seemingly similar processes, one involving conscious identification and discrimination of tonal categories (Burton, Small, & Blumstein, 2000; Zatorre, Evans, Meyer, & Gjedde, 1992), the other involving unconscious usage of tonal information for lexical access (Damasio & Damasio, 1980). According to previous studies on segmental processing, the two processes are carried out in two distinct pathways of the human brain (Hickok & Poeppel, 2000, 2007; Myers, Blumstein, Walsh, & Eliassen, 2009). In the current study, the former process is related to the tonal oddity task, and the latter to the lexical decision task.

5.2 Method

5.2.1 Participants

Twenty native speakers of Jinan, six males and fourteen females, participated in this experiment in exchange for payment. They were all right-handed and their ages ranged from 23 to 39. An oddity test was carried out first to divide the participants into two groups according to their tonal awareness (X. Chen et al., 2004; Shu et al., 2008).

5.2.2 Design and stimuli

A mixed design was adopted. The within-subject variable involved four levels of tonal variability and the between-subject variable involved two levels of tonal awareness (high and low). Disyllabic stimuli were selected from a corpus of high-frequency disyllabic Chinese words produced by six native speakers (3 male and 3 female) of Jinan Mandarin. Each word was read twice in random order. The same group of speakers also produced pseudowords. In total, we elicited four sets of 160 stimuli. All six speakers contributed to each set. The first set included twenty pairs of acoustic identity (Condition 1). The same stimulus was presented first as a prime and then as a target. In this condition, primes and targets were always identical. The second set included twenty pairs of stimuli with within-category variation (Condition 2). Within this set, each pair of stimuli showed the same tonal pattern with only very subtle differences in their actual instantiation of the pitch contour or in other acoustic dimensions (such as duration and intensity). The third set included twenty pairs of words with tonal pattern variation (Condition 3). In this condition, we picked words which were produced by the same speaker but with different tonal patterns in her/his two renditions of the same stimulus words, with one rendition as the prime and the other as target. The fourth set included lexically contrastive tonal pairs (Condition 4). In this condition, tonal minimal pairs (which share the same segmental structure) from the same speaker were chosen with one as the prime and the other as the target. In addition, 160 pseudowords and 20 real words were included as fillers. Figure 2 illustrates the four conditions with examples.

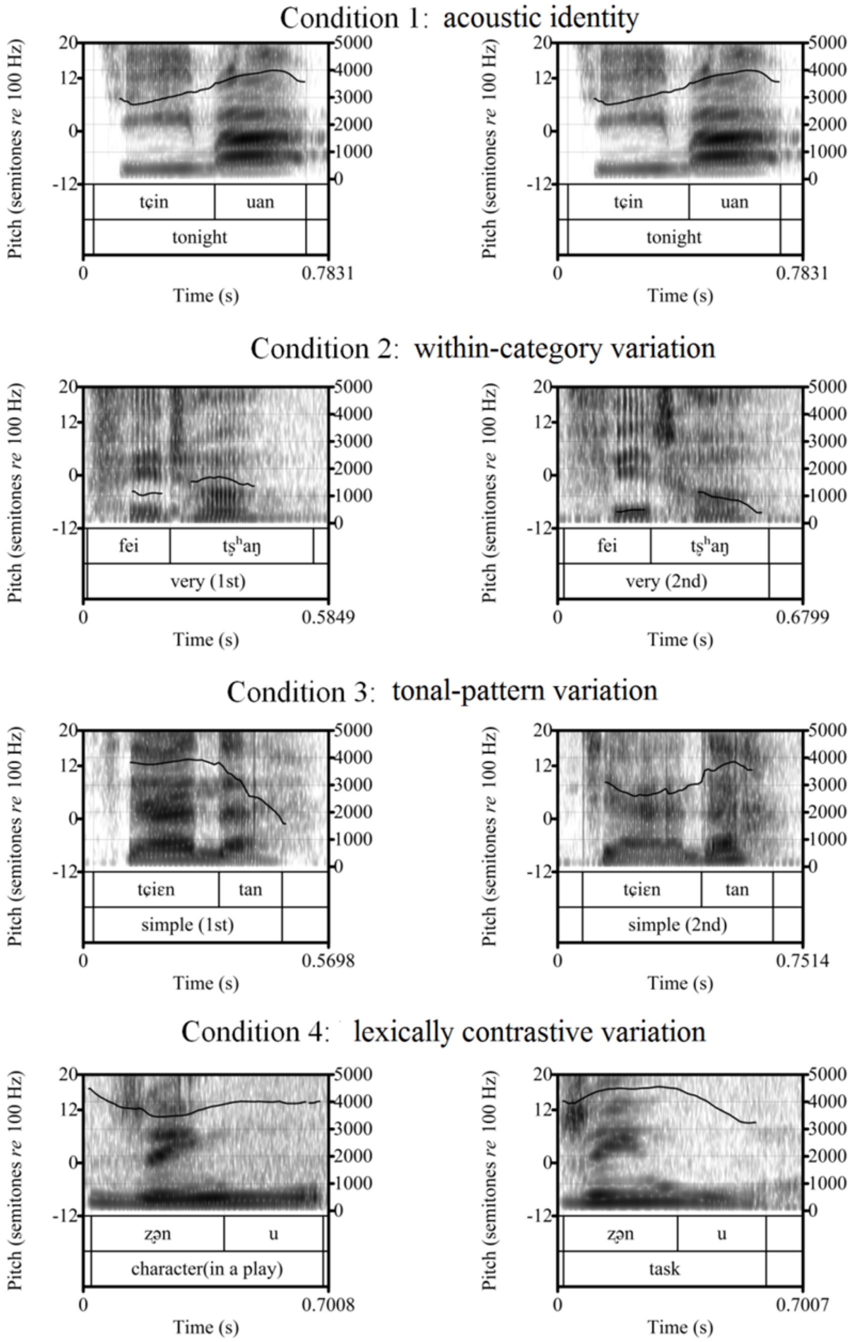


Figure 2. *Examples of the prime (left) and the target (right) in conditions 1-4, with pitch contours and spectrograms plotted.*

5.2.3 Procedure

Participants were tested individually in a quiet room. In the tonal oddity test, they listened to twelve sets of four monosyllabic Jinan words. Within each set, there was one word which had a different lexical tone, commonly labelled as the odd member. Participants pressed a number key on the keyboard to indicate the odd word (in the order of the presentation). The accuracy rate was calculated for each participant, according to which the participants were divided into two groups with high vs. low tonal awareness.

Then the participants were told that they would hear a series of disyllabic sound sequences in Jinan Mandarin and they had to decide whether or not each of these sound sequences was a real word. The presentation of the stimuli was controlled by E-Prime 2.0 run on a laptop equipped with a Creative SBX-FI5.1 pro sound card. Each item was played binaurally through headphones, with instructions on the screen. The target of each pair was played 4 to 6 items after its corresponding prime. Pairs in each group were presented in different random orders to the participants, while stimuli from the same condition were never presented directly following one another. A blank screen was displayed between every two items.

A *priming effect* was defined as the reaction time difference between the first occurrence (prime) and the second occurrence (target)(Pallier, Colomé, & Sebastián-Gallés, 2001). If the reaction to the target is faster than that to the prime, the priming effect will be positive; if it is slower, the priming effect will be negative. Here, a positive value is taken as facilitation and a negative value is taken as interference.

We performed two sets of analyses. First, Analyses of Variance (by-participants and by-items) were performed on the *priming effect* to assess the differences across tonal conditions and between participant groups. Then, mixed-linear-effect analyses were performed on the *reaction times* to further investigate the influence of tonal awareness on the process of primes and targets.

5.3 Results

5.3.1 Analysis 1

The average *priming effects* are plotted in Figure 3, as a function of the following factors: Contour Condition (acoustic identity vs. within-category variation vs. tonal pattern variation vs. lexically contrastive variation) and Tonal Awareness (low vs. high). The 95% error bars are based on the within-subjects repeated ANOVAs.

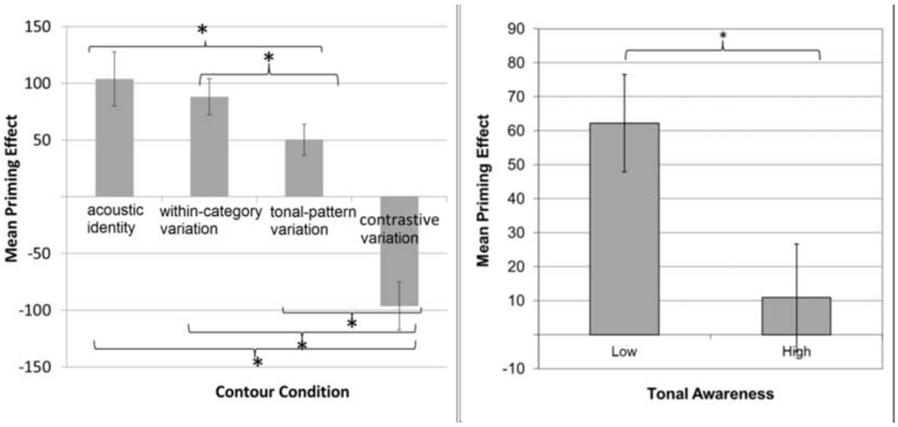


Figure 3. *Mean priming effect for each contour condition (left) and tonal awareness conditions (right).*

Contour Condition had a significant main effect (by participants, *F1*(3, 54) = 24.87, *MSE* = 6,641, $p < 0.05$, and by items, *F2*(3, 76) = 5.36, *MSE* = 28,416, $p < 0.05$). Tonal Awareness also had a significant effect in the by-participants analysis, *F*(1, 18) = 5.79, *MSE* = 9,003, $p < 0.05$, r = 0.49. There was no significant interaction between Contour Condition and Tonal Awareness. The LSD test after the by-participants analysis showed that only tonal pattern variation (Condition 3) and lexically contrastive variation (Condition 4) were significantly different from all the other conditions. Only contrastive variation (Condition 4) was significantly different from the acoustic identity and within-category variation in the Games-Howell test after the by-items analysis and in the Bonferroni test after the by-participant analysis. Generally speaking, the priming effect was smaller (lower in absolute value if positive; higher in absolute value if negative) for participants with higher tonal awareness.

The difference between cross-lexical vs. within-lexical variability was very robust. When the tonal difference between the prime and the target was lexically contrastive, the priming effect was always negative, indicating an interference effect. In other words, there seems to be an inhibition effect when word pairs with different lexical tones were presented. In the other three conditions, the difference between the prime and the target is not lexically contrastive. In this case, the priming effect was always positive, suggesting a facilitation effect. As for the within-lexical variability, the facilitation effect varied, however, according to the degree of similarity between the prime and the target. When a tonal pattern difference was involved, the priming effect was much more reduced than that of the other two conditions (i.e. acoustic identity and within-category variation conditions), where there was no significant priming difference.

5.3.2 Analysis 2

A linear mixed effects analysis was performed on the *reaction times*, using R (R_Core_Team, 2013), *lme4* (Bates, Maechler, Bolker, & Walker, 2013), and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2013). The model included the fixed effects of Condition (acoustic identity, within-category variation, tonal pattern variation, and lexically-contrastive variation), Tonal Awareness (high or low), Type of Stimulus (prime or target), and their two-way and three-way interactions, as well as the random effects of by-item intercept, and by-participant slopes for the effect of order with its intercept. Several likelihood-ratio tests were first performed to justify the way the random terms were introduced. As shown in Table 1, with Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2013; SAS, 1978), significant effects were found for items and the by-participant slopes for the effect of order. The main effect of Condition and the interactions of Tonal Awareness and Type were significant. As shown in Figure 4, Tonal Awareness had a larger effect on primes than on targets.

Figure 4. *2-way interaction plot of tonal awareness (high or low) and the type of stimuli (prime or target) on reaction time data*

Table 1. Summary of Mixed Effects Model for variables predicting reaction times (RT)

Fixed effects	Df	F	p
Condition	3	4.6549	<0.01
Tonal awareness	1	0.1023	>0.05 (n.s.)
Type	1	3.7895	>0.05 (n.s.)
Condition: Tonal awareness	3	0.2636	>0.05 (n.s.)
Condition: Type	3	1.9959	>0.05 (n.s.)
Tonal awareness: Type	1	7.6229	<0.01
Condition: Tonal awareness: Type	3	0.6909	>0.05 (n.s.)
Random effects		X^2	
1 | item	1	942	<0.0001
1 + order | participant	1	11	<0.0001

5.4 Discussion

5.4.1 Main results and interpretations

The current result shows that the two forms of tonal pattern variation share a joined representation in lexical access. Priming across these forms yielded facilitation effects, which are similar to the positive priming induced by acoustic variants but significantly different from the negative priming induced by contrastive variants. Note that the difference between the tonal pattern variation and the contrastive variation cannot be attributed to the amount of acoustic difference, because the amounts are roughly the same (see Figure 1 (b) versus 1 (c)). The only possible reason is that this priming happens at a stage where a joint representation is formed for the two tonal forms linking to the same word but the separation of two lexically contrastive forms is maintained. Thus, the interference effect may be due to the competition of the two activated lexical items, while the facilitation effect may rather be due to the previous activation of the form.

Note that the two tonal patterns in the tonal pattern variation condition are potentially contrastive in other words. This suggests that at this particular level of processing, there cannot be tonal generalization, as otherwise the minimal contrastive pairs involved would become indistinguishable as well. Our results thus support the claim that tonal pattern variation is processed in a lexically specific way.

Priming between tonal pattern variants reduced the size of the facilitation effects, compared with the priming effect induced by within-category variation. This difference cannot be attributed to any difference in lexical contrast, since it is the same in both conditions that the prime and the target are lexically non-contrastive. This difference may either be attributed to the amount of acoustic difference or to the fact that there are two tonal patterns involved in the tonal pattern variation. Although the current results cannot differentiate between these two possibilities, tonal pattern variation may also have its own representational status in the process of lexical access, different from the access of the purely acoustic tonal variability.

Moreover, the tonal awareness, which involves grouping words according to tonal categories and explicit access to tonal phonemes, also influences word recognition. The priming effect data showed that the low-tonal-awareness participants yielded more facilitation in the non-contrastive condition but less interference in the contrastive condition. The reaction time data further showed that lower tonal awareness resulted in slower lexical decisions to the primes but not to the targets. This indicates that no matter whether the corresponding lemma is selected (Conditions 1, 2, 3) or not (Condition 4), a participant's tonal awareness only affects the speed of the first activation of the lexical node. Further studies are needed to explain the mental mechanism behind tonal awareness and its contribution to lexical access.

5.4.2 Theoretical implications

According to Levelt's theory of language production, stress functions like an abstract lexical frame (Levelt, Roelofs, & Meyer, 1999). Cutler & Van Donselaar, 2001 showed that minimal stress pairs have distinctive representations in lexical access in Dutch. If tonal patterns work like stress, different tonal patterns should also have distinct representations in lexical access. Supportive evidence was found in the current study. First, contrastive tonal minimal pairs showed a different type of priming (interference) compared with non-contrastive pairs involving only one tonal pattern (facilitation). Second, the non-contrastive facilitation was significantly mitigated when different tonal patterns (though non-contrastive) were presented for the same word.

However, by including a previously untested condition, the current study also revealed new evidence for the joined representation of different tonal variants for the same word. When different tonal patterns link to the same word, the priming was positive, just like in the other lexically non-contrastive conditions. It is different from the negative priming across tonal minimal pairs. This result is in line with earlier studies on the storage of pronunciation variants, and supports that both tonal-pattern variants are stored under the same lemma (e.g. Bürki, Ernestus, & Frauenfelder, 2010; Ernestus, 2014; Pitt, 2009).

One may argue that the results are also compatible with the possibility of tonal under-specification, which leads to listeners' tolerance of phonemic tonal mismatches. This would then be similar to what have been reported on segmental mismatch and the flexibility of such mismatch in lexical processing (Connine, Blasko, & Titone, 1993; Lahiri & Marslen-Wilson, 1991; Marslen-Wilson & Zwitserlood, 1989; McClelland & Elman, 1986; McMurray, Tanenhaus, & Aslin, 2009; Milberg, Blumstein, & Dworetzky, 1988). While our results lack direct evidence to rule this possibility out, it is very important to note that our results in general cannot be explained by under-specification of lexical representation or listeners' tolerance of mismatches in processing. First, the prime and the target only varied along the tonal dimension, biasing the listeners to tune into tonal variation. Thus, it is highly unlikely that listeners were not sensitive to tonal mismatches. Second, had listeners been generally flexible about tonal information, we would not have found the different priming effects due to different types of tonal mismatch, since segmental sharing was consistent across all conditions and all conditions should have then yielded the same priming effects.

Taken together, we have shown that it is important to examine what and how different levels of tonal representations are involved in lexical access. The lexically non-contrastive tonal pattern variation, which previously received little attention in the literature, provides us with a great opportunity to look into the role of tonal patterns in lexical access. The current results support that different tonal variants have separate representations at the lexeme level and share the same representation under the same lemma. The evidence for the joined representation is relatively new.

Some questions require further investigation. When does the convergence of lexically non-contrastive tonal patterns take place in the process? Is it processed at a separate stage? Note that the convergence needs to be lexically specific because otherwise other words distinguished via these tonal patterns would be confused. How to incorporate the observed convergence in the models of lexical access remains an open question.

This study provides evidence for the representational status of tonal patterns in lexical access, while the existence of joint tonal representations depends on the existence of lexical contrasts. The effect of tonal awareness is mitigated when a similar form has been heard shortly before. On aggregate, these results call for a model of lexical access, which incorporates the representations of tonal patterns, the lexical specific convergence of non-contrastive tonal forms, and a mechanism to account for its interaction with tonal awareness.

Acknowledgments

We would like to thank Prof. Shengli Cheng, Prof. Xiufang Du, Jia Li, Lulu Zhou, and Dianliu Neighbourhood Committee for the recruitment of participants. J. Wu's work was supported by a PhD Studentship sponsored by Talent and Training China-Netherlands Program. Y. Chen is currently supported by the European Research Council (ERC Starting Grant-206198). The fieldtrip was sponsored by the Leiden University Foundation and Leiden University Centre for Linguistics.

References

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition, 52*(3), 163-187. doi: 10.1016/0010-0277(94)90042-6

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one 'fenêtre' in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language, 62*(4), 421-437. doi: 10.1016/j.jml.2010.01.002

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: an fMRI investigation. *Journal of Cognitive Neuroscience, 12*(4), 679-690. doi: 10.1162/089892900562309

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751-781. doi: 10.1006/jmla.2001.2825

Chen, X., Anderson, R. C., Li, W., Hao, M., Wu, X., & Shu, H. (2004). Phonological Awareness of Bilingual and Monolingual Chinese Children. *Journal of Educational Psychology, 96*(1), 142-151. doi: 10.1037/0022-0663.96.1.142s

Chen, Y., Shen, R., & Schiller, N. O. (2011). *Representation of allophonic tone sandhi variants.* Paper presented at the Proceedings of Psycholinguistics Representation of Tone. Satellite Workshop to ICPhS, Hongkong.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic bulletin & review, 11*(6), 1084-1089. doi: 10.3758/BF03196741

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*(2), 193-210. doi: 10.1006/jmla.1993.1011

Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics, 59*(2), 165-179. doi: 10.3758/BF03211886

Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and speech, 44*(2), 171-195. doi: 10.1177/00238309010440020301

Damasio, H., & Damasio, A. R. (1980). The anatomical basis of conduction aphasia. *Brain, 103*(2), 337-350. doi: 10.1093/brain/103.2.337

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*(1), 27-41. doi: 10.1016/j.lingua.2012.12.006

Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech, 1*(2), 126-154.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*(4), 131-138. doi: 10.1016/S1364-6613(00)01463-7

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience, 8*(5), 393-402. doi: 10.1038/nrn2113

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://CRAN.R-project.org/package=lmerTest

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245-294. doi: 10.1016/0010-0277(91)90008-R

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *BEHAVIORAL AND BRAIN SCIENCES, 22*(1), 1-75. doi: 10.1017/S0140525X99001776

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 576-585. doi: 10.1037/0096-1523.15.3.576

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86. doi: 10.1016/0010-0285(86)90015-0

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from ‘lexical’ garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language, 60*(1), 65-91. doi: 10.1016/j.jml.2008.07.002

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126. doi: 10.1207/s15516709cog0000_79

Milberg, W., Blumstein, S., & Dworetzky, B. (1988). Phonological processing and lexical access in aphasia. *Brain and Language, 34*(2), 279-293. doi: 10.1016/0093-934X(88)90139-3

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Attention, Perception, & Psychophysics, 65*(6), 956-969. doi: 10.3758/BF03194826

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: evidence from a learning paradigm. *Cognitive Science, 35*(1), 184-197. doi: 10.1111/j.1551-6709.2010.01140.x

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics, 34*(1), 73-103. doi: 10.1016/j.wocn.2005.03.003

Myers, E. B., Blumstein, S. E., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science, 20*(7), 895-903. doi: 10.1111/j.1467-9280.2009.02380.x

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449. doi: 10.1111/1467-9280.00383

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61*(1), 19-36. doi: 10.1016/j.jml.2009.02.005

Qian, Z.-Y. (1995). Jinan Fangyan Cidian Yinlun (Introduction to the 'Jinan Dialect Dictionary'). *Fangyan [Dialects], 95*, 242-256.

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

SAS. (1978). Tests of Hypotheses in Fixed-Effects Linear Models *SAS Technical Report R-101* Cary, NC, USA: SAS Institute Inc.

Shu, H., Peng, H., & McBride - Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11*(1), 171-181. doi: 10.1111/j.1467-7687.2007.00654.x

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Yip, M. J. W. (2002). *Tone*. Cambridge: Cambridge University Press.

Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*(5058), 846-849. doi: 10.1126/science.256.5058.846

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhou, X., Q, Y., Shu, H., Gaskell, G., & Marslen-Wilson, W. (2004). Constraints of lexical tone on semantic activation in Chinese spoken word recognition. *Acta Psychologica Sinica, 36*(4), 379-392.

6 Do Tonal Bilinguals Store the Unproduced Tonal Variant of Etymologically Related Words? [10]

Abstract

Tonal bilinguals of Jinan Mandarin and Standard Chinese (SC) produce different tonal variants of the same Jinan word. Usually all the variants are segmentally identical to the word's SC counterpart but only one of the variants is tonally identical to the word's SC counterpart (variant_id). The word-wise probability of variant_id varies between 0 and 1. Naming latency data were elicited for 400 Jinan words from 42 speakers to test whether speakers who produced the variant_id also store the unproduced variant which is not tonally identical to the SC counterpart (variant_ni). If the speakers who produced the variant_id do not store the unproduced variant_ni, the naming latency should only depend on the speaker's choice of variant_id (yes/ no) but not on the word-wise probability of variant_id. If the speakers who produced the variant_id do store the unproduced variant_ni, the naming latency should depend on the word-wise probability of the speaker's chosen variant. The latter was verified by our results.

6.1 Introduction

Etymologically related words are translation equivalents which are similar in sound. They are either inherited from the common ancestor as cognates or borrowed across languages as loan words.

The phonological similarity between a pair of etymologically related words varies along a continuum and etymologically related words can be practically distinguished as identical and non-identical. For instance, the Dutch 'computer' and the English 'computer' are more similar than the Dutch 'neus' and the English 'nose'. Experimental evidences suggest that etymologically related words with different degrees of phonological similarity may have different statuses in lexical representation and lexical access. For instance, only identical cognates showed cognate facilitation effect in eye-tracked reading, while non-identical cognates did not (Duyck, Assche, Drieghe, & Hartsuiker, 2007). However, the effect of phonological similarity is inconsistent within and across studies (Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Duyck et al., 2007; Lemhöfer & Dijkstra, 2004). Several reasons may be responsible for the unstable effects. First, the phonological similarity co-varies and interacts with the orthographic similarity. Second, etymologically related words can be non-identical in different ways. For instance, some have different vowels and the others have different consonants. However, in the bilingualism considered by these studies, there are not enough cases for studying each sub-type separately.

---

[10] This chapter is based on Wu, J., & Chen, Y. (2014). Tonal variants in the bilingual mental lexicon. Paper presented at the The Fourth International Symposium on Tonal Aspects of Languages (TAL 2014), Nijmegen. ISBN: 978-90-9028606-8

The degree of phonological similarity becomes a more important variable when it comes to the tonal bilingualism of Jinan and standard Chinese (SC). First, unlike in the previous studies, the etymologically related words are the majority in the vocabulary of bilinguals who speak these two closely related dialects. As a result, the majority of translation equivalents are phonetically similar, providing much more cases for studying this phenomenon systematically. Second, both Jinan and standard Chinese are written with the same logographic Chinese writing system. Thus all Jinan-SC etymologically related words are orthographically identical, which controls the confusion from orthography. Third, the segmental differences between the etymologically related words are almost reduced to annihilation in the youngest generation. The tonal similarity between a Jinan word and its counterpart in standard Chinese decides the phonological similarity between them. As a result, the two etymologically related words are only different in tone. We can focus on the role of tone in the bilingual lexical representation and access.

Moreover, some Jinan words show different variants identical and non-identical to their SC counterparts in speech production. In our Jinan corpus collected in 2012, some Jinan words were produced with two or more tonal patterns, as shown in Figure 1. Here we call them multi-pattern words. Such a word usually has a variant almost identical to the word's counterpart in SC (variant_id), together with one or more variant(s) which is/are not identical to the words SC counterparts (variant_ni). Note that the segmental structure of both variant_id and variant_ni are almost always identical to the multi-pattern words' counterparts in SC and the only difference is carried by tone.

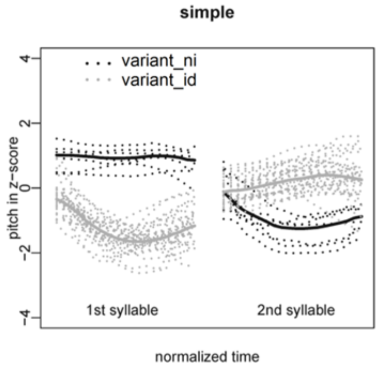


Figure 1: *The Illustration of one multi-pattern word. It is made with recordings produced by 42 Jinan speakers (1 or 2 outliers excluded). Pitch values were z-transformed semitones (the mean and SD were calculated with about 600 recordings for each speaker).*

How the two variants are stored in different speakers' memory is still an open question. It is related with the following two factors. First, for one Jinan word, a speaker can produce either a variant_id or a variant_ni in one rendition. Second, with each speaker providing one rendition of each word, the probability of variant_id can be measured for each word, by calculating the percentage of speakers who produced variant_id for the word. The former measurement represents whether the variant_id of a lexical item is retrieved in the actual rendition of speech production (speaker's choice of variant_id, yes/ no); the latter represents the general probability of variant_id when the bilingual speakers name the shared orthographic form of the pair of etymologically related words in Jinan (word-wise probability of variant_id). For instance, the SC word ‘simple’ realizes with a low-plus-high tone. If we observed Speaker A produced the Jinan word ‘simple’ with a low-plus-high tone (variant_id) in his rendition and Speaker B produced the Jinan word ‘simple’ with a high-plus-rising tone (variant_ni) in his rendition, we say Speaker A made the choice of variant_id (y) and Speaker B did not make the choice of variant_id (n). With 32 out of 41 speakers producing the same word ‘simple’ in Jinan and, the word-wise probability of variant_id for the Jinan ‘simple’ is 0.78.

As shown in Figure 2, word-wise probability of variant_id ranges between 0 and 1 in our corpus. The majority of words were only produced with Jinan-only variants. The other 123 of the 400 (25.5%) recorded words were produced identical to its SC counterpart by at least one speaker. Within these words, 33 were produced identical to its SC counterpart by more than 85% of the speakers and 21 by around half (31%-76%) of the speakers. It is unlikely that all or most of the speakers coincidentally make the same code-mixing error together. This phenomenon indicates that some variant_id should be natively Jinan and tagged as a Jinan lexical item in the mental lexicon. On the other hand, the majority (66) of the multi-pattern words were produced only by a few (2%-26%) speakers as identical to its SC counterpart, which could be the real examples of code-mixing.

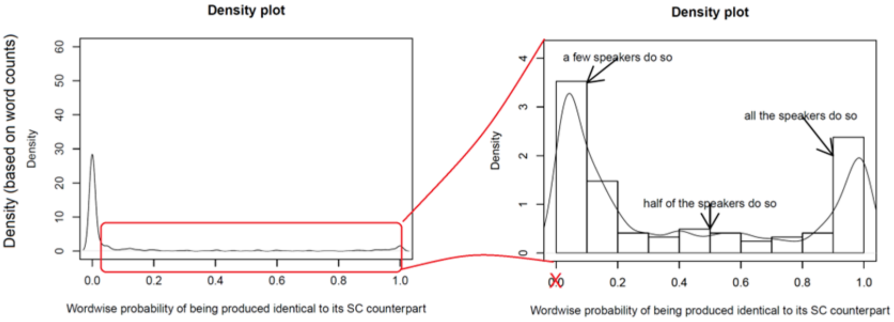


Figure 2: *Density plots of the word-wise probability of variant_id.*

Do speakers who did not produce variant_id also store it in an integrated lexicon together with variant_ni? The answer to this question decides how we should interpret the effect of word-wise probability of variant_id theoretically. The status of variant_id in the bilingual mental lexicon yields different predictions for the following experimental questions. Does the word-wise probability of variant_id affect the naming latency of the word, no matter which variant is picked in the actually rendition? Does the speaker's choice of variant_id affect the naming latency of the word, no matter how low or high the word-wise probability of variant_id is for this word? Do the two predictors interact?

Assuming the lexical representations made up by different phonemes (including tone) are also different (Dijkstra et al., 1999), variant_id and variant_ni should have different lexical representations. As shown in Figure 3, (1) one possibility is that different variants are stored by different speakers. The speakers who produced variant_ni store variant_ni and the others who produced variant_id store variant_id as the phonological representation for the same Jinan word. The former speakers have both variant_ni and variant_id and the latter speakers only have variant_id. Under this hypothesis, the word-wise probability of variant_id only reflects the distribution of individual difference of phonological representation in the language system and should not affect the naming latency of the multi-pattern word. Instead, the speaker's choice of variant_id should show effects. Assuming an integrated bilingual lexicon (Van Heuven, Dijkstra, & Grainger, 1998), speaker who produced variant_ni should be generally slower than speakers who produced variant_id because the former speaker's variant_ni (Jinan) receives extra competition from variant_id.

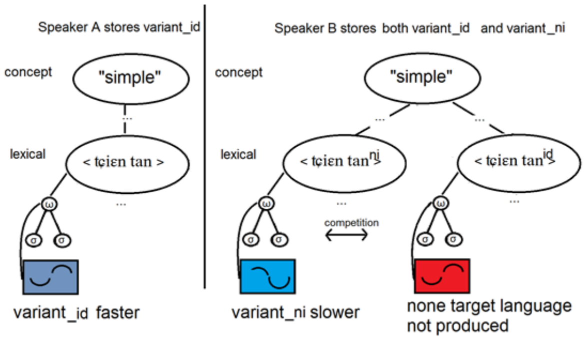


Figure 3: *Possibility (1). Different speakers store different variants.*

As shown in Figure 4 the other possibility is that all the speakers store both the variant_id and the variant_ni in an integrated lexicon. Under this hypothesis, the word-wise probability of variant_id reflects the likelihood of variant_id being selected in Jinan lexical access. In this case, the more likely the produced variant is, the shorter the naming latency should become, and no matter it is variant_id or variant_ni. We expect a higher word-wise probability of variant_id should reduce the naming latency of variant_id and increase the naming latency of variant_ni, since the condition implies a relatively lower likelihood of variant_ni. Correspondingly, a lower word-wise probability of variant_id should increase the naming latency of variant_id and reduce the naming latency of variant_ni, since the condition implies a relatively higher word-wise probability of variant_ni. Thus an interaction of speaker's choice of variant_id and word-wise probability of variant_id should be observed.

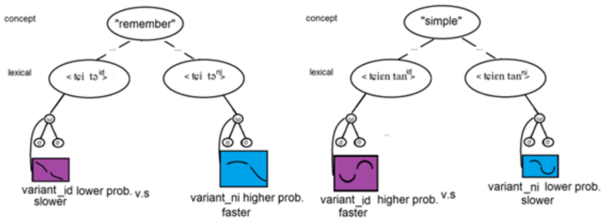


Figure 4: *Possibility (2). The variant with higher word-wise probability is produced faster.*

6.2 Experiment

6.2.1 Data preparation

Participant. The speech data used in the present study were collected from 42 Jinan native speakers in 2012.

Stimuli. Each speaker read 400 disyllabic Chinese words in Jinan. The written words were selected from a corpus of Chinese film subtitles (Cai & Brysbaert, 2010), in a way that one list of 200 high-frequency words were selected from the 10% disyllabic Chinese words with the highest word frequency. In a similar way, we selected the other list of 200 low-frequency words. In each list, each of the 20 disyllabic tonal combinations of standard Chinese contributes 10 words.

Procedure. The high and low frequency lists were presented to the speakers in two blocks with a self-paced rest break in between. The words in each list were presented in a different random order for each speaker. After the speaker finished producing a word, they pressed a key to see the next word.

A trained phonetician listened to each recording, looked at the spectrogram, and manually marked the beginning and the rhyme of the production. Also In this process, recordings with speech and recording errors were excluded from the corpus. Then naming latencies and pitch contours on the rhymes were extracted. To further remove pitch contour outliers, we calculated Local Outlier Factors (LOF) for each speaker's z-normalized pitch contours on the rhyme. Any pitch contours with an LOF greater than 1.5 (Breunig, Kriegel, Ng, & Sander, 2000) and belong to the 2.5% with the highest integral density were eliminated from the corpus. The naming latency outliers were excluded using a (method I) distributional based approach (van der Loo, 2010) on the log transformed naming latency.

Whether the Jinan word was produced almost identical to its counterpart in standard Chinese was judged by a phonetician with Putonghua Proficiency Test Certificates-Level1B. The probability of variant_id can be measured for each word, by calculating the percentage of speakers who produced variant_id for the word. The word-frequency was grouped into two levels (high-low) using the Chinese subtitle data (Cai & Brysbaert, 2010).

6.2.2 Model fitting

Only renditions of multi-pattern words ($N = 3368$) were taken into consideration in the following analysis. Linear mixed effects analyses were performed on the naming latency data, using R (R_Core_Team, 2013), lme4 (Bates, Maechler, Bolker, & Walker, 2013), and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2013) in the following way. A full model was built first, including the fixed effects of the speaker's choice of variant_id, the word-wise probability of variant_id, the Chinese word frequency, the tonal categories of the words' SC counterpart, and their two-way and three-way interactions, as well as the random intercept of the word and the speaker. (The random intercept model was proven to be better than alternative random slope models via model comparisons). When there were unrealized combinations of the nominal predictors, the corresponding interaction terms were removed. A backward elimination was then performed to remove non-significant effects, using p-values calculated from F test based on Sattethwaite's method (Kuznetsova et al., 2013).

In the final model, the main effect of Chinese word frequency was the only significant main effect, $F = 15.61$, $p < 0.05$. A higher Chinese word frequency reduces the Jinan naming latency. The main effect of the speaker's choice of variant_id, $F = 0.92$, n.s., and the word-wise probability of variant_id, $F = 0.10$, n.s., were both insignificant. However, their interaction was significant, $F = 4.00$, $p < 0.05$. As shown in Figure 5, for a word with higher word-wise probability of variant_id, the variant_id was named faster than the variant_ni; for a word with lower word-wise probability of variant_id, the variant_id was named slower than the variant_ni. All the other fixed terms were insignificant and removed in the model trimming.

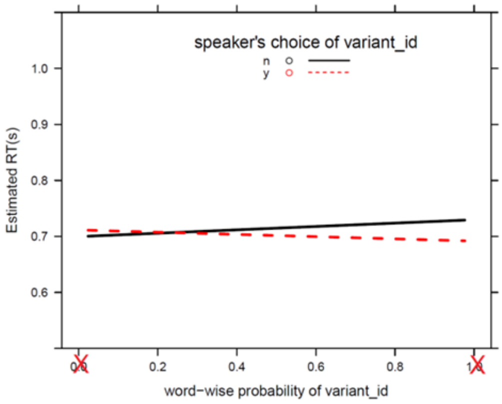


Figure 5: *Interaction of the speaker's choice of variant_id and the word-wise probability of variant_id.*

6.3 Discussion

The results support the hypothesis that all the speakers store both the variant_id and the variant_ni in an integrated lexicon (Figure 2.2), which divide the effect of Chinese word frequency. The result is also in-line with the data from a smaller corpus where each word was produced twice by each speaker in different random orders. We have observed both variant_ni and variant_id in the two renditions by the same speaker. We also have shown that the priming between such two variants spoken by the same speaker is similar (but reduced) compared with the priming between two renditions of the same variant in median-term auditory priming using lexical decision task (Wu, Chen, Schiller, & Van Heuven, accepted).

On the other hand, we have also seen in another study that individual backgrounds affect the tonal realizations of variant_ni. This indicates that individual differences have their own impact on the phonological representation (Wu et.al in preparation). However, considering this effect with the finding of the current experiment, individual differences seem to have their effect more on the shape of the stored tonal patterns than the lexical storage of variants.

Some variant_id are very likely native Jinan and stored in the Jinan lexicon. To include variant_ids in the Jinan lexicon, the theory should either allow (a) duplicated lexical nodes or (b) duplicated tagging for this variant.

6.4 Conclusions

The Jinan-SC tonal bilinguals' naming latencies of Jinan words depended on the word-wise probability of the speaker's chosen variants. No matter the chosen variant is tonally identical or non-identical to the word's SC counterpart, the higher the word-wise probability of the variant is, the shorter the naming latency is. The result supports that the speaker who produced the variant which is tonally identical to the words' SC counterpart do store the unproduced variant which is not tonally identical to the words' SC counterpart.

Acknowledgements

We would like to thank Prof. Shengli Cheng, Prof. Xiufang Du, Jia Li, Lulu Zhou, and Dianliu Neighbourhood Committee for the recruitment of participants. J.Wu's work was supported by a PhD Studentship sponsored by Talent and Training China-Netherlands Program.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). *LOF: identifying density-based local outliers.* Paper presented at the ACM Sigmod Record.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301.

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 663.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package).

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32*(4), 533-550.

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

van der Loo, M. P. J. (2010). Distribution based outlier detection for univariate data,. *Discussion paper 10003, Statistics Netherlands, The Hague*.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language, 39*(3), 458-483.

Wu, J., Chen, Y., Schiller, N. O., & Van Heuven, V. J. (accepted). Tonal Variability in Lexical Access. *Language and Cognitive Processes*.

7 One Writing System, Two Dialects: Tonal Information in Bilingual Visual Word Recognition

Abstract

How is the tonal information of written words activated when the reader is a bilingual of two tonal dialects? Standard Chinese (SC) and Jinan Mandarin (JM) are two tonal dialects and share a large number of etymologically related translation equivalents, which are represented by the same Chinese character and usually identical in segments but potentially different in tone. In the present study, we adopted a Stroop paradigm to examine the activation of tonal information in Chinese visual word recognition by tonal bilinguals and tonal monolinguals. Native Beijing SC monolinguals and Jinan JM-SC bilinguals named the ink color of Chinese characters in SC and the bilinguals also performed the color-naming in JM. The Chinese characters were either (1) color characters (e.g. 红, hong2, 'red'), (2) within-dialect homophones of the color characters (S+T+; e.g. 洪, hong2, 'flood'), (3) within-dialect different-tone homophones of the color characters (S+T−; e.g. 轰 , hong1, 'boom'), or (4) neutral characters (S−T−; e.g. 贯, guan, 'penetrate') as controls. We tested all the combinations of characters and ink colors. Both groups of participants showed classic Stroop interference for the incongruent conditions. However, only the bilinguals showed Stroop facilitation for the color-congruent conditions. They are slower in production but better at handling conflicts in lexical access. The within-dialect tonal effects were more complex than those found in earlier studies. Only the tonal monolinguals, not the tonal bilinguals, showed weak evidence in support of the retrieval of tonal information in automatic visual word recognition. Between-dialect tonal relations showed no additional effects on tonal bilinguals.

7.1 Introduction

Bilingual visual word recognition has been investigated in depth in earlier studies. It is clear that the bilinguals activate relevant lexical items from both languages and the interlingual activation happens even when the two languages use different types of orthographies. For instance, bilingual Stroop effects have been verified in many studies, where the color naming was in one language and the words were printed in another language, either with the same type of orthography or not (Chen & Ho, 1986; Dyer, 1971; Fang, Tzeng, & Alva, 1981; Kiyak, 1982; Preston & Lambert, 1969). A recent trilingual study further showed that the between-language interference and facilitation are both affected by the similarity of the scripts (Van Heuven, Conklin, Coderre, Guo, & Dijkstra, 2011).All these studies showed robust between-language interference while the interference is usually smaller between languages than within languages (MacLeod, 1991).

However, little is known about the automatic lexical activation by bilinguals who use the same logographic writing system for different languages or dialects, especially with regard to phonological activation. The bilingualism of related Chinese dialects involves a large number of etymologically related translation equivalents (i.e. cognates), which are similar in sound and written with the same Chinese characters. In such cases, one Chinese character can be associated with two similar but distinctive sounds in the mental lexicon of the bilinguals. Moreover, in the bilingualism involving Standard Chinese with a northern Mandarin dialect, such as Jinan Mandarin (JM) (Wu & Chen, 2014), the two interlingual pronunciations for the same Chinese character are usually almost identical in segments but can vary in their tonal similarity. This tonal bilingualism in related dialects is a common phenomenon in China but a special case of bilingualism.

Tonal bilingualism was not taken into consideration in earlier studies on Chinese visual word recognition. For instance, previous Chinese Stroop experiments did not report whether Standard Chinese (Mandarin) is their participants' only Chinese dialect (Li, Lin, Wang, & Jiang, 2013; Spinks, Liu, Perfetti, & Tan, 2000). It may therefore be the case that tonal bilinguals may read Chinese characters in a different way from monolinguals. If so, many questions remain open to answers. First, are tonal bilinguals different from tonal monolinguals in their automatic visual word recognition? Second, which phonological representations are activated in the bilingual mental lexicon by the common Chinese character? Are the different tonal representations from both dialects automatically activated via the same Chinese character?

Stroop-related paradigms have long been used to investigate visual word recognition (MacLeod, 1991). The classical Stroop effect (Stroop, 1935) emerges when participants are shown different words written in different colors. When they are asked to name the word, the ink color has no influence on the word naming time. However, when they are asked to name the ink color, they are unable to suppress the influence from the word. When the word and the ink color are congruent (e.g. 'RED' in red ink), the word will facilitate the color naming relative to an unrelated control word (Dalrymple-Alford, 1972). When the word and the ink color are incongruent (e.g. 'RED' in green ink), the word will interfere with the color naming relative to different types of controls, such as color patches, 'X's, unrelated words, and non-words (MacLeod, 1991). Without specific training (MacLeod & Dunbar, 1988; Stroop, 1935), the ink colors are usually ignored in word naming but the words cannot be ignored in color naming.

The Stroop effect is sensitive to the phonological relations between the printed word and the color name. Earlier studies on alphabetic writing systems verified that when the printed word shares phonemic features with a color word which is incongruent with the ink color, the interference increases (Dalrymple-Alford, 1972) (Dennis & Newstead, 1981; Singer, Lappin, & Moore, 1975; Underwood & Briggs, 1984), and when the printed word shares phonemic features with the color name, the interference is reduced, or the color naming is facilitated, depending on whether color patches or irrelevant words are used as the control condition (Effler, 1978). The effect size increases with the amount of phonemic sharing (Dennis & Newstead, 1981; MacLeod, 1991) and shows an initial letter effect (Dalrymple-Alford, 1972; Dennis & Newstead, 1981; Singer et al., 1975; Underwood & Briggs, 1984). Moreover, comparing pseudo-homophones of color words (e.g. *bloo* for *blue*) with non-words matched for visual similarity and initial phonemic overlap with color words (e.g. *blir* for *blue*) showed that the phonemic effects in alphabetic writing systems are not just a result of orthographic similarity (Dennis & Newstead, 1981).

The phonological effects are neither limited to alphabetic writing systems nor limited to segmental relations. Similar phonemic effects have also been observed in Chinese, which uses a logographic (sometimes also termed as 'ideographic') writing system and has lexical tones. Both segmental and tonal sharing between the printed word and the color character (the orthographic unit that represents a syllable with a specific meaning and tone) showed the phonemic effects, in that segmental homophones of color characters facilitated the naming of congruent colors and interfered with the naming of incongruent colors (Li et al., 2013; Spinks et al., 2000), and additional tonal sharing slightly added to the facilitation from segmental homophones in two of the experiments (Spinks et al., 2000, Experiment 1 and Li et al., 2013). Also, Li et al. (2013) found phonological facilitation based on tonal sharing alone. Using a related picture-word interference paradigm, it has also been found that the distractor and target matching in underlying tonal category and overt tonal realization facilitated picture naming (Nixon, Chen, & Schiller, 2014). Compared with the results using alphabetic scripts, the Stroop interference with Chinese characters is at least of the same strength (Lee & Chan, 2000) (Smith & Kirsner, 1982) if not greater (Biederman & Tsao, 1979; Saalbach & Stern, 2004; Tsao, Wu, & Feustel, 1981), although it has been shown that Chinese Stroop tasks may involve more right hemisphere interference (Tsao et al., 1981). The Stroop phonemic effects in Chinese are also in line with the more general findings, with segmental information carrying more weight than tonal information (Li et al., 2013; Spinks et al., 2000).

The phonological effects found in Stroop paradigms reflect the automatic activation of phonological information in visual word recognition. Although there are different theoretical accounts for the Stroop effect (Cohen, Dunbar, & McClelland, 1990; MacLeod, 1991; Posner & Snyder, 1975), it is now generally agreed that, when the task is color naming, the activation of the printed words is relatively automatic and can escape the attention to some extent (Cohen et al., 1990), differently from when the task is word naming. This automatic phonological activation is also valid in Chinese Stroop paradigms, although the Chinese writing system is logographic. In previous studies (Li et al., 2013; Spinks et al., 2000), the pronunciations of the Chinese characters affected the color naming latency. Together with other findings using the priming paradigm (Perfetti & Zhang, 1991, 1995), the Stroop phonological effects in Chinese indicate that the phonological information carried by the Chinese characters is automatically activated even if the task does not require word naming.

Moreover, compared with the other Stroop-like paradigms (such as the picture-word interference paradigm), the classical color-word paradigm holds a relatively more constant attention demand for the non-word dimension across trials. This allows more detailed manipulations on the word dimension. These features make the Stroop paradigm especially interesting for investigating the automatic activation of phonological information by bilinguals.

In the present study, the phonological effects in Stroop paradigm can be used to differentiate the activations of JM and SC lexical nodes and tonal representations in the automatic visual word recognition. If we find phonological sharing increases facilitation or reduces interference, the corresponding phonological information should be activated. The phonological relation between the printed color character and the ink color can be manipulated within and between dialects, yielding different levels and dimensions of phonological sharing. The different color-character combinations can also serve as controls for one another.

We replicated the Stroop experiment (Li et al., 2013; Spinks et al., 2000) on Jinan tonal bilinguals in Standard Chinese (SC) and Jinan Mandarin (JM), as well as on Beijing tonal monolinguals in SC. Tonal bilinguals and monolinguals both named the ink color of Chinese characters in SC and the bilinguals also named the ink color in JM. Taking the between-dialect conditions into consideration, additional stimuli were added and some stimuli were replaced to avoid unintended between-dialect sharing. As shown in Table1.1, considering the tonal relation of the character and its related color characters within each dialect, the Chinese characters were either

(1) color characters (e.g. 红, hong2, ‘red’),
(2) within-dialect homophones of the color characters (S+T+; e.g. 洪, hong2, ‘flood’),
(3) within-dialect different-tone homophones of the color characters (S+T−; e.g. 轰, hong1, ‘boom’), or
(4) neutral characters (S−T−; e.g. 贯, guan, ‘penetrate’).

Each within-dialect condition except the neutral characters was shown with both congruent and incongruent ink colors (e.g. 红 ‘red’, 洪 ‘flood’, and 轰 ‘boom’ in red versus in green). Differently from the previous studies (Li et al., 2013; Spinks et al., 2000), which only tested each character with one congruent and one incongruent color, we tested all the combinations of characters and ink colors. With the reaction times of the neutral trials for the same colors as the base-lines, we expect to find Stroop facilitation for the congruent trials and Stroop interference for the incongruent trials, as was found in earlier studies (Li et al., 2013; Spinks et al., 2000). The corresponding statistics on the effects of Stroop congruence are provided in Analysis 1 and the different within-dialect phonological conditions were compared in Analysis 2.

Two comparisons were made to replicate the S−T+ (tone-only) condition tested in Li et al.’s (2013) study in Analysis 2, 3, and 4. First, as shown in Table 1.2, the neutral (S−T−) characters designed for some colors shared the tone with another color (S−T+) [e.g. 贯, guan2, ‘penetrate’, is S−T− for 红 ‘red’ but share the same tone (S−T+) for ‘绿’ lǜ2] in both dialects[11].Such tonal congruent and tonal incongruent neutral-character trials were tested against each other for every ink color. Li et al.’s (2013) finding predicts a significant tone-only facilitation of S−T+ trials compared with S−T− trials. Second, as shown in Table 1.3, a S+T− trial presented with an incongruent ink color can share the tone with the name of the ink color or not (e.g. 览, lan3, ‘view’, S+T− with 蓝, lan2, ‘blue’, is incongruent with

---

[11] Note that due to the tonal systematic correspondence across the two dialects, characters from the same tonal category in SC are mostly also from the same tonal category in JM and we only included the stimuli which follow the systematic corresponding rules. For instance, here the tonal marking ‘2’ is referring to both the JM high-falling and SC high-rising tones.

both green and purple ink colors but only shares the tone with the name of purple, zi3). If the tonal information is independently activated in the automatic visual word recognition, such S+T− trials with incongruent ink colors of the same tone (e.g. 览 in purple) should show reduced Stroop interference compared with the other incongruent S+T− trials (e.g. 览 in green).

The effects of the three language modes, SC-monolingual, SC-bilingual, JM-bilingual and their interactions with the within-language condition were examined for the bilingual effects. Bilinguals have been found to be slower in lexical retrieval (Bialystok, 2009; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ransdell & Fischler, 1987; Rogers, Lister, Febo, Besing, & Abrams, 2006) but better at solving the conflict of tasks (Bialystok, 2009; Carlson & Meltzoff, 2008; Hilchey & Klein, 2011; Prior & Gollan, 2011), including the conflict of the Stroop task (Bialystok, Craik, & Luk, 2008). However, it remains unclear whether earlier findings on bilinguals of non-tonal languages are applicable to tonal bilinguals of related dialects. If tonal bilinguals are like other bilinguals, we expect the tonal bilinguals to be generally slower than monolinguals in the Stroop task but show less Stroop interference. Moreover, Spinks et al. (2000) and Li et al. (2013) did not specify whether their Mandarin (SC) speakers are tonal monolinguals or bilinguals. The present study will clarify which of their findings are replicable on both native bilinguals and monolinguals and which are only true for one of the speaker groups.

In the present study, we also investigated between-dialect tonal activation by comparing trials which have the same within-dialect tonal relation in SC but different between-dialect tonal relations, as shown in Table 1.4 and Table 1.5. For instance, both 烂, lan4, 'rotten' and 览, lan3, 'view' are S+T− to the color character 蓝, lan2, 'blue' in SC. However, different from 'view', the SC pronunciation of 'rotten' carries a falling tone (Falling), which turns the SC version of 'rotten' to a false friend of the JM pronunciation of 'blue' (JM 'blue' lan2 = /lan(High-falling)/). As another example, 绿, lǜ2, 'green', with falling tones in both dialects, is an identical cognate in JM and SC, but the other color characters carry different tonal contours in JM and SC. If there is cross-dialect tonal activation in automatic visual word recognition, beside the general bilingual effects, we should observe additional differences in regard to these special characters between the two groups of participants (in Analysis 5 and 6).

Table 1.1 Critical trials for the within-Dialect conditions

Condition	Color Characters	S+T+	S−T−	S+T−[a]	S−T− (neutral)
Character	蓝	栏	烂	览	抱
Translation	blue	fence	rotten	view	hold
Frequency [b]	102.59	32.37	48.50	52.81	131.29
N strokes	13	9	9	9	8
Pronunciation SC	lan(Hr)	lan(Hr)	lan(F)	lan(Lr)	pau(F)
Pronunciation JM	lan(Hf)	lan(Hf)	lan(Lf)	lan(Hl)	pau(Lf)
Common pinyin	lan2	lan2	lan4	lan3	bao4
Character	绿	虑	旅/驴		涂
Translation	green	consider	travel/ donkey		smear
Frequency	133.39	180.17	153.50/ 26.45		59.16
N strokes	11	10	10/ 6		10
Pronunciation SC	ly(F)	ly(F)	ly(Lr)/ ly(Hr)		tʰu(Hr)
Pronunciation JM	ly(Lf)	ly(Lf)	ly(Hl)/ ly(Hf)		tʰu(Hf)
Common pinyin	lü4	lü4	lü3/ lü2		tu2
Character	紫	子	自		鼻
Translation	purple	child	self		nose
Frequency	99.84	4001.12	3113.46		83.87
N strokes	12	3	6		13
Pronunciation SC	tsɿ(Lr)	tsɿ(Lr)	tsɿ(F)		pi(Hr)
Pronunciation JM	tsɿ(Hl)	tsɿ(Hl)	tsɿ(Lf)		pi(Hf)
Common pinyin	zi3	zi3	zi4		bi2
Character	红	洪	轰		贯
Translation	red	flood	boom		penetrate
Frequency	419.08	122.34	79.11		89.37
N strokes	6	9	8		8
Pronunciation SC	xuŋ(Hr)	xuŋ(Hr)	xuŋ(Hl)		kuan(F)
Pronunciation JM	xuŋ(Hf)	xuŋ(Hf)	xuŋ(R)		kuan(Lf)
Common pinyin	hong2	hong2	hong1		guan4
Character	黄	皇	荒	谎	岸
Translation	yellow	emperor	shortage	lie	shore
Frequency	478.04	438.15	106.25	27.59	16.46
N strokes	11	9	9	11	8
Pronunciation SC	xuaŋ(Hr)	xuaŋ(Hr)	xuaŋ(Hl)	xuaŋ(Lr)	an(F)
Pronunciation JM	xuaŋ(Hf)	xuaŋ(Hf)	xuaŋ(R)	xuaŋ(Hl)	an(Lf)
Common pinyin	huang2	huang2	huang1	huang3	an4

a. JM High-falling is undergoing merging with JM High-level

b. Character per million frequencies on the Chinese Text Computing website (Da, 2004; http://lingua.mtsu.edu/chinese-computing/).

c. Abbreviation for tones. SC: Hl = High-level (1), Hr = High-rising (2), Lr = Low-rising (dip tone) (3), F = Falling (4). JM: R=Rising (1), Hf=High-falling (2), Hl=High-level (3), Lf=Low-falling (4)

Table 1.2 S−T− and S−T+ trials combined with different ink colors

Color		S−T+	S−T−
Blue Red Yellow (SC: Hr JM: Hf)	Character	涂/鼻	抱/贯/岸
	Translation	smear/nose	hold/penetrate/shore
	Frequency	59.16/83.87	131.29/89.37
	Number of strokes	10/13	8/8/8
	Pronunciation SC	tʰu(Hr)/pi(Hr)	pau(F)/kuan(F)/an(F)
	Pronunciation JM	tʰu(Hf)/pi(Hf)	pau(Lf)/kuan(Lf)/an(Lf)
	Common pinyin	tu2/bi2	bao4/guan4/an4
Green (SC: F JM: Lf)	Character	抱/贯/岸	涂/鼻
	Translation	hold/penetrate/shore	smear/nose
	Frequency	131.29/89.37	59.16/83.87
	Number of strokes	8/8/8	10/13
	Pronunciation SC	pau(F)/kuan(F)/an(F)	tʰu(Hr)/pi(Hr)
	Pronunciation JM	pau(Lf)/kuan(Lf)/an(Lf)	tʰu(Hf)/pi(Hf)
	Common pinyin	bao4/guan4/an4	tu2/bi2

Table 1.3 S+T− trials combined with T− or T+ incongruent ink colors.

S+T− characters		Related color	T+ color	T− color
Character	览			
Translation	view	blue	purple	green/red/yellow
Pronunciation SC	**lan(Lr)**	**lan**(Hr)	tsɿ(**Lr**)	ly(F)/xuŋ(Hr)/xuaŋ(Hr)
Pronunciation JM	**lan(Hl)**	**lan**(Hf)	tsɿ(**Hl**)	ly(Lf)/xuŋ(Hf)/xuaŋ(Hf)
Common pinyin	**lan3**	**lan**2	zi**3**	lü4/hong2/huang3
Character	旅			
Translation	travel	green	purple	blue/red/yellow
Pronunciation SC	**ly(Lr)**	**ly**(F)	tsɿ(**Lr**)	lan(Hr)/xuŋ(Hr)/xuaŋ(Hr)
Pronunciation JM	**ly(Hl)**	**ly**(Lf)	tsɿ(**Hl**)	lan(Hf)/xuŋ(Hf)/xuaŋ(Hf)
Common pinyin	**lü3**	**lü**4	zi**3**	lan2/hong2/huang2
Character	谎			
Translation	lie	yellow	purple	blue/green/red
Pronunciation SC	**xuaŋ(Lr)**	**xuaŋ**(Hr)	tsɿ(**Lr**)	lan(Hr)/ly(F)/xuŋ(Hr)
Pronunciation JM	**xuaŋ(Hl)**	**xuaŋ**(Hf)	tsɿ(**Hl**)	lan(Hf)/ly(Lf)/xuŋ(Hf)
Common pinyin	**huang3**	**huang**2	zi**3**	lan2/lü4/hong2

Table 1.4 Test pairs for the between-dialect tonal relation (1): within-dialect S+T− to the same corresponding color characters, S+T+ v.s. S+T− between-dialect false friends.

	S+T+ between dialects	S+T− between dialects
Character	烂	览
Translation	rotten	view
Frequency [b]	48.50	52.81
N. strokes	9	9
Pronunciation SC	lan(F)	lan(Lr)
Pronunciation JM	lan(Lf)	lan(Hl)
Common pinyin	lan4	lan3
Within-dialect condition	S+T− to the ColChar. ‘blue’	S+T− to the ColChar. ‘blue’
Notes	SC lan(F) (rotten) ≈ JM lan(Hf) (blue)	
Character	驴	旅
Translation	donkey	travel
Frequency [b]	26.45	153.50
N. strokes	6	10
Pronunciation SC	ly(Hr)	ly(Lr)
Pronunciation JM	ly(Hf)	ly(Hl)
Common pinyin	lü2	lü3
Within-dialect condition	S+T− to the ColChar. ‘green’	S+T− to the ColChar. ‘green’
Notes	JM ly(Hf) (donkey) ≈ SC ly(F) (green)	
Character	荒	谎
Translation	shortage	lie
Frequency [b]	106.25	27.59
N. strokes	9	11
Pronunciation SC	xuaŋ(Hl)	xuaŋ(Lr)
Pronunciation JM	xuaŋ(R)	xuaŋ(Hl)
Common pinyin	huang1	huang3
Within-dialect condition	S+T− to the ColChar. ‘yellow’	S+T− to the ColChar. ‘yellow’
Notes	SC xuaŋ(Hl)(shortage) ≈ JM xuaŋ(Hf) (yellow)	

ColChar. = Color Character; N. strokes = Number of Strokes

Table 1.5 Test sets for the between-dialect tonal relation (2): color characters, as S+T+ v.s. S+T− cognates between dialects, and S+T+ homophones to the corresponding color characters, as S+T+ v.s. S+T− false friends between dialects.

Between-Dial. color cognates	S+T+	S+T−			
Character	绿	蓝	紫	红	黄
Translation	green	blue	purple	red	yellow
Frequency [b]	133.39	102.59	99.84	419.08	478.04
N. strokes	11	13	12	6	11
Pronun. SC	ly(F)	lan(Hr)	tsɿ(Lr)	xuŋ(Hr)	xuaŋ(Hr)
Pronun. JM	ly(Lf)	lan(Hf)	tsɿ(Hl)	xuŋ(Hf)	xuaŋ(Hf)
Common pinyin	lü4	lan2	zi3	hong2	huang3
Within-dialect condition	ColChar. ‘green’	ColChar. ‘blue’	ColChar. ‘purple’	ColChar. ‘red’	ColChar. ‘yellow’
Notes	JM ly(Lf) (green) ≈ SC ly(F) (green)				
Between-Dial. false friends	**S+T+**	**S+T−**			
Character	虑	栏	子	洪	皇
Translation	consider	fence	child	flood	emperor
Frequency [b]	180.17	32.37	4001.12	122.34	438.15
N. strokes	10	9	3	9	9
Pronun. SC	ly(F)	lan(Hr)	tsɿ(Lr)	xuŋ(Hr)	xuaŋ(Hr)
Pronun. JM	ly(Lf)	lan(Hf)	tsɿ(Hl)	xuŋ(Hf)	xuaŋ(Hf)
Common pinyin	lü4	lan2	zi3	hong2	huang3
Within-dialect condition	S+T+ to the ColChar. ‘green’	S+T+ to the ColChar. ‘blue’	S+T+ to the ColChar. ‘purple’	S+T+ to the ColChar. ‘red’	S+T+ to the ColChar.’ yellow’
Notes	JM ly(Lf) (consider) ≈ SC ly(F) (green)				

ColChar. = Color Character; N. strokes = Number of Strokes

7.2 Experiment

7.2.1 Participants

Forty-eight native tonal monolinguals of SC from Beijing (7 male and 41 female, aged between 19 and 30, M = 22.73, SD =2.95) and 54 native SC-JM tonal bilinguals from Jinan (16 male and 38 female, the age ranged from 19 to 36, M = 22.70, SD = 3.85, 44 SC dominant or balanced, 10 JM dominant) participated in this experiment in exchange for payment. Both groups are right-handed, received their literacy educations in SC, and have learned some English in school. Four participants from the Jinan group and 15 participants from the Beijing group also have some knowledge of other non-tonal foreign languages, such as French and German.

7.2.2 Design and stimuli

An unbalanced mixed design was adopted. The stimuli were presented in five different ink colors (blue, green, purple, red, and yellow) including 230 (23 characters × five colors × two repetitions) critical trials and ten neutral training trials. Within the 23 characters, four were related to the color character of ‘blue’, four were related to the color character of ‘green’, three were related to the color character of ‘purple’, three were related to the color character of ‘red’, four were related to the color character of ‘yellow’, and the remaining five were neutral characters each assigned to one color character. The five or four characters related to the same color character include the color character, the S+T+ character, the S+T− character, JM merging S+T− character (when available), and neutral character (see Table 1.1for the stimulus characteristics). Each non-neutral character was congruent with one color and incongruent with the remaining four colors. Additional within-dialect and between-dialect relations were considered separately (see Table 1.2, 1.3, and 1.4 for details) in Analyses 2 to 6. Beijing monolinguals of SC were tested with all the stimuli in SC mode and Jinan bilinguals were tested with the same stimuli in both SC and JM modes.

7.2.3 Procedure

The experiment was implemented using the E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Participants named the color of the characters shown on the screen as quickly and accurately as possible. Each trial started with the presentation of a fixation cross that appeared in the center of the screen for 1,000 ms, followed by the target character printed in 48-point SimSun font, which disappeared after 2,000 ms. Then the following trial started. A recording was made from the appearance of each target character until the appearance of the next character. Critical trials were preceded by ten neutral training trials. The two repetitions of the critical trials were split into two blocks and the trials within each block appeared in different randomized orders for each participant. The order of the trials in the whole experiment was recorded for further analysis. Half of the bilinguals were first tested in the SC mode and then in the JM mode and the other half were first tested in the JM mode and then in the SC mode. The language mode was prompted by the dialect of the auditory instruction and five auditory examples of color naming. The bilinguals had a short break and other tasks (first in the previously tested language mode and then in the coming language mode) between the two Stroop experiments in different modes to avoid abrupt switching. The monolinguals were tested with exactly the same procedure (with auditory instruction and other tasks before) but only in the SC mode.

The naming latencies of the recordings were automatically measured with a Praat (Boersma & Weenink, 2014) script (Pacilly, 2010) based on the intensity of the sound pressure. Then a trained phonetician (the first author) listened to each recording, looked at the waveform and spectrogram, and manually corrected any errors in the marking. The naming accuracy of each recording was also manually marked in this process.

7.2.4 Analysis and discussion

Six sets of linear mixed-effect (LME) models were built to investigate the Stroop effects (Analysis 1), within-dialect segmental effects (Analysis 2), within-dialect tonal effects (Analysis 2, 3, 4), and between-dialect tonal effects (Analysis 5, 6). Reaction time (RT) analysis was based on the correct trials only. To improve the distribution of the data, RT data were log-transformed. Naming latency outliers were excluded for each participant using a distribution based approach (van der Loo, 2010), method I) on the log transformed naming latency. *Trial Order in the Same Color* and *Trial Distance from the Same Color* were calculated for each trial using the trial order data. Table 2 shows descriptive statistics for RTs and error rates organized by within-dialect condition and language modes. The reaction times in the congruent conditions were subtracted from reaction times in the neutral conditions for facilitation (negative) and interference (positive) effects.

Since the design was unbalanced, we performed linear mixed-effect (LME) analysis on different subsets of the log-transformed reaction times to investigate the influences and interactions of different between- and within-participant predictors. The analyses were performed using R (R_Core_Team, 2013), lme4 (Bates, Maechler, Bolker, & Walker, 2013), and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2013). The LME models are summarized in the Appendix. The significance was calculated with Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2013; SAS, 1978) but the more transparent numeric degrees of freedom were reported. Separate models were used to analyze congruent and incongruent trials. Random terms include by-item and by-participant intercepts or slopes for the effect of *Trial*, *Trial Distance from the Same Color*, and/or *Trial Order in the Same Color,* selected via model comparison based on likelihood ratio tests. The post-hoc contrasts between conditions were calculated with the lmerTest package and the significance was marked in Table 2, 3.1 and 3.2. We performed six sets of LME analyses as shown in the following part.

7.2.5 Results

Statistics of all the LME models are summarized in the Appendices. The rest of the results are presented as follows.

Analysis 1 Stroop Effects. To investigate the classical Stroop congruence effects, the first LME analysis compared each non-neutral within-dialect condition with the corresponding neutral characters. In these models, *Within-Dialect Condition* (a non-neutral condition versus the corresponding neutral condition), *Mode*, *Target Color*, *Trial Order in the Same Color*, *Trial Distance from the Same Color,* and their interactions were the fixed predictors.

The main effects of *Within-Dialect Condition* were significant in all the models. Compared with the neutral condition, significant facilitations emerged for all the congruent color characters and significant interference emerged for all the incongruent color characters. The main effects of *Mode* were significant in all the models, except the model comparing JM tonal-merging characters with the neutral characters presented with congruent colors. The post-hoc analysis of Differences of Least Squares Means (DLSM) showed that bilinguals named colors generally more slowly than monolinguals but the difference between the bilinguals' SC and JM modes was insignificant. The interaction between *Within-Dialect Condition* and *Mode* was significant in all the models, except the two models comparing the (S+T−) and JM tonal-merging characters with the neutral characters presented with congruent colors. The post-hoc analysis of DLSM showed that the congruent conditions facilitated the bilinguals more than the monolinguals and the incongruent conditions affected the bilinguals less than the monolinguals. Different from the bilinguals, the monolinguals showed no significant Stroop facilitation but only Stroop interference, as shown in Table 2.

The Stroop interference and facilitation is consistent with the earlier findings (Li et al., 2013) (Spinks et al., 2000) and indicate that phonological information in Chinese is activated automatically in visual word recognition. Tonal bilinguals are slower than tonal monolinguals in the Stroop task but they benefit more from the congruent conditions and suffered less from the incongruent conditions. This is consistent with the other earlier findings that bilinguals are slower in lexical retrieval (Bialystok, 2009; Gollan et al., 2005; Ransdell & Fischler, 1987; Rogers et al., 2006) but better at solving conflicts of tasks (Bialystok, 2009; Bialystok et al., 2008; Carlson & Meltzoff, 2008; Hilchey & Klein, 2011; Prior & Gollan, 2011). The Stroop interference of both bilinguals and monolinguals was consistent with Spinks et al.'s (2000) findings. The Stroop facilitation of the tonal bilinguals was also consistent with the earlier findings (Li et al., 2013; Spinks et al., 2000). However, the tonal monolinguals showed no such Stroop facilitation. Since the previous studies did not specify whether their Mandarin speakers had a background in other Chinese dialects, the Stroop facilitations found by Spinks et al. (2000) and Li et al. (2013) may be specifically due to tonal bilingualism instead of the usage of tone.

Table 2 Response times of correct trials (RTs, with SDs in parentheses) and errors (with SDs in parentheses) in each within-dialect condition

	JM			SC-bi			SC-mono		
Condition	RT (ms)	Errors (%)	Stroop Effects	RT (ms)	Errors (%)	Stroop Effects	RT (ms)	Errors (%)	Stroop Effects
Cong. ColChar.	706 (193)	0.2 (1.4)	-20 ***	688 (177)	0.4 (2.0)	-38 ***	700 (193)	0.0 (0.0)	19 ns.
Cong. S+T+	686 (176)	0.0 (0.0)	-40 ***	692 (173)	0.6 (2.4)	-34 ***	683 (174)	0.2 (1.4)	2 ns.
Cong. S+T−	701 (187)	0.7 (2.3)	-25 ***	708 (172)	0.2 (1.2)	-18 **	688 (172)	0.7 (2.3)	7 ns.
Cong. S+T− JM merging	685 (169)	0.0 (0.0)	-41 ***	673 (162)	0.0 (0.0)	-53 ***	674 (157)	0.0 (0.0)	-7 ns.
InC. ColChar.	821 (227)	3.5 (5.4)	95 ***	810 (220)	2.6 (3.2)	84 ***	828 (220)	3.0 (3.2)	147 ***
InC. S+T+	757 (195)	1.7 (3.0)	31**	752 (187)	1.2 (2.0)	26**	741 (192)	1.7 (2.8)	60***
InC. S+T−	752 (194)	1.2 (2.3)	26 **	744 (185)	0.7 (1.1)	18 *	730 (183)	0.9 (1.5)	49 ***
InC. S+T− JM merging	746 (205)	1.0 (2.3)	20 **	736 (183)	1.1 (3.0)	10 ns.	722 (174)	0.5 (1.7)	41 ***
S−T− (neutral controls)	726 (199)	0.4 (2.0)		726 (199)	0 (0)		681 (157)	0.4 (2.9)	

ns. $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$.
Cong. = Congruent; ColChar. = Color Character; InC. = Incongruent.

Analysis 2 Segmental and (Lack of) Tonal Effects. To investigate the within-dialect phonological Stroop effects, the second LME analysis excluded neutral trials and compared different within-dialect conditions listed in Table 1.1, with congruent and incongruent conditions analyzed in separate models. This analysis sheds lights on the different roles of segmental and tonal information in visual word recognition. In these models, *Within-Dialect Condition* (color characters/homophones of the color characters/different-tone homophones of the color characters/different-tone homophones of the color characters with potential tonal merging in JM), *Mode*, *Target Color*, *Trial*, *Trial Distance from the Same Color*, and their interactions were the fixed predictors.

The main effect of *Within-Dialect Condition* was significant in the model for the incongruent conditions but insignificant in the model for the congruent conditions. The main effect of Mode was significant in both models, showing similar bilingual disadvantage of lexical access as in Analysis 1. The interaction of *Within-Dialect Condition* and *Mode* was significant in the model for the incongruent conditions but insignificant in the model for the congruent conditions. The descriptive statistics of each within-dialect condition are shown in Table2. In the post-hoc analysis of DLSM, the Stroop interference from color characters was greater than that from the three types of homophones. However, whether the homophone shared the same tone with the corresponding color character did not affect the strength of Stroop interference, except that the bilinguals in SC mode showed greater interference with S+T+ homophones than with S+T− homophones with potential tonal merging in JM. The post-hoc analysis on the model for congruent trials showed few significant contrasts, with only two exceptions. First, when bilinguals performed in SC mode, the Stroop facilitation was significantly greater for the color characters than for the S+T− homophones. Second, when bilinguals performed in the JM mode the Stroop facilitation was significantly smaller for the color characters than for the S+T+ homophones.

The stronger Stroop interference from color characters is consistent with the earlier study by Spinks (2000) and can be attributed to the semantic and lexical activation of the characters. However, the stronger Stroop facilitation from color characters (Li et al., 2013; Spinks et al., 2000) was only partially replicated by the bilinguals but not at all by the monolinguals, indicating that the Stroop facilitation found earlier may be specific to Chinese tonal bilinguals. Actually, as was also shown in Analysis 1, monolinguals showed no significant Stroop facilitation. As for the Stroop facilitation with the bilinguals, the pattern was also inconsistent in different modes. Only the bilinguals' Stroop facilitations in SC mode were similar to earlier findings (Li et al., 2013; Spinks et al., 2000). Instead, in JM mode it is the S+T+ homophones that showed the greatest Stroop facilitation. The different strength of facilitation between S+T+ and S+T− homophones found by Li (2013) and Spinks (2000, Exp. 1) was not replicated in any mode or either group of participants. Thus although the results of Analysis 1 support the activation of phonological information of the characters, no evidence was found in Analysis 2 in support of additional tonal activation beside the activation of the segmental structure.

Analysis 3 & 4 Tonal effects. The third and fourth LME analysis further investigated the effect of tonal information within the target dialect. The data were stratified according to the target colors in the third analysis. A set of LME models specifically compared S−T+ and S−T− trials for each target color (as shown in Table 1.2, e.g. 贯, guan4, 'penetrate', is S−T−, but 鼻, bi2, 'nose', is S−T+ for the target color red, hong2). In these models, *Same Tone* (the character is S−T− / S−T+ with the ink color), *Mode*, *Trial*, *Trial Distance from the Same Color*, and their interactions were the fixed predictors. Similarly, in the fourth analysis the trials shown in Table 1.3 were stratified according to the language mode (SC-bilingual, JM-bilingual, & SC-monolingual). A set of LME models specifically compared T− and T+ incongruent ink colors combined with S+T− characters (e.g. 览, lan3, 'view', S+T− with the color character 蓝, lan2, 'blue', is S−T− for the target color green, lǜ2, but S−T+for the target color purple, zi3) in each mode. In these models, *Color* (T+ color / T− color), *Within-Dialect Condition* (S+T− character / neutral character), *Trial*, *Trial Distance from the Same Color*, and their interactions were the fixed predictors.

In Analysis 3 the main effects and interactions of *Same Tone* and *Mode* were insignificant with most ink colors. As shown by the descriptive statistics and the significance of the post hoc DLSM contrasts in Table 3.1, compared with common neutral (S−T−) characters, characters sharing the tone alone (S−T+) with target ink color did not facilitate color naming. Instead when the ink color was yellow, the ink color was produced significantly slower with S−T+ characters than with S−T− characters in SC modes.

Table 3.1 Response times of correct trials (RTs, with SDs in parentheses) in S−T− and S−T+ trials combined with each color

	JM			SC-bi			SC-mono		
	S−T−	S−T+[a]	Effect T+	S−T−	S−T+	Effect T+	S−T−	S−T+	Effect T+
blue	769 (175)	789 (179)	20 ns.	760 (186)	795 (190)	35 ns.	738 (170)	761 (191)	23 ns.
green	752 (222)	742 (192)	-10 ns.	773 (187)	757 (180)	-16 ns.	722 (168)	745 (176)	23 ns.
red	641 (164)	658 (162)	17 ns.	653 (162)	653 (162)	0 ns.	624 (169)	642 (149)	18.
yellow	677 (188)	695 (197)	18 ns.	661 (183)	689 (186)	28*	640 (152)	682 (193)	42*

ns. $p > .05$; * $p < .05$; . $p < 0.1$

a. The S−T+ characters are ‘hold’, ‘penetrate’, and ‘shore’ for green and ‘smear’ and ‘nose’ for the other colors. The S+T− characters for green and the other colors are the reverse.

These results are in contrast with the tone-alone Stroop facilitations found by Lin et al. (2013). In the present study the tone-alone effect was generally absent and when it is significant, the effect was interference instead of facilitation. The difference could be attributed to the fact that we used different S−T+ characters. However, our findings at least suggest that the tone-alone effect in automatic visual word recognition is not robust.

Nevertheless, we did find new evidence in support of tonal activation in Analysis 4. The crucial trials in Analysis 4 were homophone characters which shared the segmental structure but not the tone (S+T−) with the relevant color character and shared the tone but not the segmental structure (S−T+) with the ink color (e.g. 览, lan3, ‘view’, S+T− with 蓝, lan2, ‘blue’, is S−T+ with the ink color name, zi3, ‘purple’).The deviation of these crucial trials from the neutral trials with the same ink color (e.g. the deviation of 览, lan3, ‘view’ in the T+ color, purple, from the neutral character 鼻, bi2, ‘nose’ in purple) was taken and the size of interference and compared with the deviation of the other incongruent trials (with the same S+T− character as in the crucial trials) from corresponding neutral trials (e.g. the deviation of 览, lan3, ‘view’ in a T− color, green, from the neutral character 涂, tu2, ‘smear’ in green). As expected, the main effects of *Color* were significant in all the models. However, since the ink colors are generally different in naming latencies (as shown in Analysis 1), the main effect of *Color* is not necessarily due to the difference in tonal sharing. On the other hand, the main effects of *Within-Dialect Condition* (S+T− character / neutral character) were only significant in the models for the monolinguals' data and the models for the character 谎, huang3, 'lie'. As was also shown earlier in Analysis 1, the bilinguals generally received less interference in the incongruent trials. Correspondingly in the models of Analysis 4, the contrasts between S+T− characters and neutral characters did not reach significance for the bilinguals. The interactions between *Color* and *Within-Dialect Condition* were also insignificant except in the model for the monolinguals' color naming with the character 谎 as the distracter. Crucially, post-hoc DLSM contrasts (in Table 3.2 and Figure 1), however, showed that the monolinguals reduced the Stroop interference for the crucial trials (S+T− characters combined with S−T+ ink colors), except when the related color or ink color was yellow.

Thus, when the S+T− trials shared the tone with the incongruent ink color [e.g. 览, lan3, 'view', related to blue, in purple (zi3)], the monolinguals showed reduced interference. This finding is in support of tonal activation in automatic visual word recognition. However, this tonal effect was on the one hand conditioned by the related and target colors and on the other hand restricted to monolinguals. Taken together with the absence of tonal effects in Analyses 2 and 3, only tonal monolinguals but not tonal bilinguals may retrieve some tonal information in automatic visual word recognition.

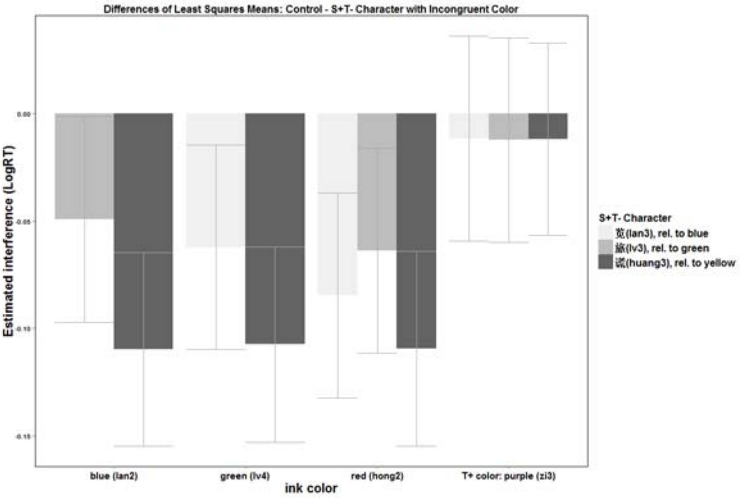


Figure 1. *Estimated interference in log-transformed RT, measured as the DLSM contrasts between the estimated RTs of the neutral trials and the S+T− incongruent trials for the same ink color, e.g. 涂, tu2, 'smear' in green (neutral trial) vs. 览, lan3, 'view', S+T− to 蓝, lan2, 'blue', in green (S+T− incongruent trial). The grey-scale filling of the bars represents S+T− characters and the bars are clustered according to the ink color. Note that negative values indicate that the trial of interest is slower than the corresponding neutral trial, and the size of the negative bars indicates the size of interference.*

Table 3.2 Additional Tone-specific effects of incongruent S+T− trials (mean RTs of the corresponding neutral controls in parenthesis)

Incongruent ink colors		green	**purple**	red	yellow
		(S−T−)	**(S−T+)**	(S−T−)	(S−T−)
Character	览	JM:	JM:	JM:	JM:
Translation	view	753	786	695	689
Pronunciation SC	lan(Lr)	(749)	(770)	(664)	(699)
Pronunciation JM	lan(Hl)	SC-bi:	SC-bi:	SC-bi:	SC-bi:
Common pinyin	lan3	770	751*	648	666
Related Color	blue	(768)	(789)	(638)	(671)
Within-Dialect	S+T−	SC-mono:	SC-mono:	SC-mono:	SC-mono:
Between-Dialect		740 *	736	657*	671
		(699)	(734)	(598)	(648)
Incongruent ink colors		blue	**purple**	red	yellow
		(S−T−)	**(S−T+)**	(S−T−)	(S−T−)
Character	旅	JM:	JM:	JM:	JM:
Translation	travel	804*	793	676	711.88
Pronunciation SC	ly(Lr)	(749)	(770)	(664)	(699)
Pronunciation JM	ly(Hl)	SC-bi:	SC-bi:	SC-bi:	SC-bi:
Common pinyin	lü3	778	772	665	675
Related Color	green	(766)	(789)	(638)	(671)
Within-Dialect	S+T−	SC-mono:	SC-mono:	SC-mono:	SC-mono:
Between-Dialect		768*	747	637*	683
		(728)	(734)	(598)	(648)
Incongruent ink colors		green	**purple**	red	blue
		(S−T−)	**(S−T+)**	(S−T−)	(S?T−)
Character	谎	JM:	JM:	JM:	JM:
Translation	lie	765*	784	696	804*
Pronunciation SC	xuaŋ(Lr)	(749)	(770)	(664)	(749)
Pronunciation JM	xuaŋ(Hl)	SC-bi:	SC-bi:	SC-bi:	SC-bi:
Common pinyin	huang3	791	786	689*	787
Related Color	yellow	(768)	(789)	(638)	(766)
Within-Dialect	S+T−	SC-mono:	SC-mono:	SC-mono:	SC-mono:
Between-Dialect		769*	742	652*	814*
		(699)	(734)	(598)	(728)

* mean RT is different from the mean RT of the corresponding neutral control ($p < 0.05$);

Analysis 5 & 6. The fifth and sixth LME analyses aimed to investigate between-dialect tonal activation. The fifth LME analysis tested the S+T+ between-dialect false friends against their S+T− counterparts (as shown in Table 1.4) considering their interaction with the dialect mode. In these models, *Between-Dialect Condition* (S+T+ false friend of the corresponding color character/S+T− false friend of the corresponding color character), *Mode*, *Trial*, *Trial Distance from the Same Color*, and their interactions were the fixed predictors.

The main effects of *Between-Dialect Condition* and *Mode* were significant in the models for the incongruent trials but insignificant in the models for the congruent trials. The interactions of *Between-Dialect Condition* and *Mode* were insignificant in all the models. This means any differences across between-dialect conditions by bilinguals were also shown in the models for monolinguals. Thus the differences found across different between-dialect conditions were not due to bilingualism but other factors, possibly the inherent differences between the specific characters.

The sixth LME analysis tested the color character, which is an S+T+ cognate or an S+T+ false friend against the other between-dialect S+T− counterparts (as shown in Table 1.5) considering their interaction with the dialect mode. In these models, *Between-Dialect Condition* (S+T+ cognate or false friend of the corresponding color character/S+T− cognate or false friend of the corresponding color character) *Mode*, *Trial*, *Trial Distance from the Same Color*, and their interactions were the fixed predictors.

The main effects of *Between-Dialect Condition* were significant in the model for the congruent cognate and the models for false-friend trials. The main effects of *Mode* were significant in the models for the incongruent trials but insignificant in the models for the congruent trials. However, similar to what was found in Analysis 5, the interactions of *Between-Dialect Condition* and *Mode* were insignificant in all the models, which means the significance of *Between-Dialect Condition* cannot be attributed to bilingualism. Taken together, the between-dialect tonal relations had no additional effects on tonal bilinguals.

7.3 General discussion

Tonal bilinguals are different from tonal monolinguals in automatic visual word recognition. First, tonal bilinguals named colors generally slower than monolinguals (as shown in Analyses 1 & 2 and Table 2), which is consistent with previous findings of bilingual lexical disadvantage on bilingual productions (Bialystok, 2009; Martin et al., 2012; Ransdell & Fischler, 1987). However, another study by Wu and colleagues (Wu, Chen, Van Heuven, & Schiller, in prep)[12] instead found lexical advantage in the same groups of participants in auditory lexical decision. This suggests that the tonal bilinguals' lexical disadvantage may be restricted to production, unlike the bilinguals using different languages.

Second, tonal bilinguals are less sensitive to tonal conditions. Only tonal monolinguals but not tonal bilinguals showed weak evidence in support of retrieval of tonal information (as in Analysis 4 and Table 3.2). Also, between-dialect tonal

---

[12] Chapter 4 in this thesis

relations showed no additional effects with tonal bilinguals compared with tonal monolinguals (as in Analyses 5 and 6). The tonal bilinguals' lack of sensitivity to tonal relations appears inconsistent with previous findings (Li et al., 2013; Nixon et al., 2014), which did find tonal effects in Stroop and Stroop-like paradigms. Moreover, our previous studies on JM-SC bilinguals' lexical decision (Wu, Chen, Van Heuven, & Schiller, 2014) and word naming (Wu & Chen, 2014) also suggest that the tonal bilinguals do make use of tonal information to distinguish tonal minimal pairs and lexical variants. However, most previous findings either did not take tonal bilingualism into consideration or adopted other paradigms. Note that the visual word recognition in the Stroop paradigm is largely automatic (Cohen et al., 1990) and the SC-JM pronunciations of the same character, if different, are only different in tone. It is reasonable then, to consider the possibility that these tonal bilinguals are better at redirecting their attention away from the Chinese characters while naming the ink color and maintain a shallower processing of the Chinese characters when it is beneficial. As shown in earlier studies, tonal information is retrieved more slowly than segmental information (Ye & Connine, 1999; Zhang & Damian, 2009). The tonal bilinguals might have already completed the color naming before the visual word decoding reaches the tonal information of the character. Thus the tonal bilinguals' lack of tonal sensitivity in the Stroop paradigm may actually reflect their fine-tuned attention control, which benefits their handling of Stroop interference.

Indeed, we found more evidence in support of the tonal bilinguals' advantage in handling the conflicts of tasks (as in Analysis 1 and Table 2). On the one hand, tonal bilinguals benefited more from the congruent conditions and suffered less from the incongruent conditions. On the other hand, only tonal bilinguals enjoyed Stroop facilitation but tonal monolinguals did not. The findings in the two previous studies which did not distinguish tonal bilinguals from tonal monolinguals (Li et al., 2013; Spinks et al., 2000) are mostly consistent with the tonal bilinguals' responses in SC mode in the current study. Hence we suppose Spinks et al. (2000) and (Li et al., 2013) might have actually used tonal bilinguals. It is not surprising to see the tonal bilinguals' advantage in handling the character-color incongruence, considering the well-known bilingual advantage in conflict resolution (Bialystok, 2009; Carlson & Meltzoff, 2008; Hilchey & Klein, 2011; Prior & Gollan, 2011).Similar bilingual advantages were also found in previous studies using Stroop tasks (Bialystok et al., 2008), although they used a different type of control condition.

It is even more interesting to see that only tonal bilinguals but not tonal monolinguals responded faster to congruent trials than to neutral-character trials (as shown in Analysis 1 and Table 2). The tonal monolinguals' lack of Stroop facilitation is very different from what Spinks et al. (2000) and Li et al. (2013) found and this cannot be simply attributed to the difference in the speakers' knowledge of English, as a foreign language. All the participants in the two previous studies and the current study learned English at school, and the proficiency can be ranked from high to low as follows: Chinese student in the US (Li et al., 2013; Spinks et al., 2000), Beijing SC tonal monolinguals, and JM-SC tonal bilinguals. If English proficiency were the cause of the inconsistency in Stroop facilitation, we would have seen a more similar pattern between the tonal monolinguals and the Chinese students in the US, which is not what we found. Nevertheless, another explanation may be more reasonable. Besides the advantages in conflict solving, bilinguals have also shown advantages in using alerting cues (Costa, Hernández, Costa-Faidella, & Sebastián, 2009; Costa, Hernández, & Sebastián-Gallés, 2008). The visual word forms are processed faster than ink colors in the Stroop paradigm (MacLeod, 1991). Thus the congruent characters can serve as alerting cues for the target colors compared with irrelevant characters. The tonal bilinguals presumably make better use of these cues than the tonal monolinguals. Note that the tonal monolinguals are also late Chinese-English bilinguals but this bilingualism did not bring them any facilitation in these congruent trials. Thus if the advantage we found is a bilingual alerting advantage, this alerting advantage should be sensitive to the status of the cue in their bilingual situation. The tonal bilinguals and tonal monolinguals are not different in whether they know more than one language. However, the tonal bilinguals use Chinese characters for both of their dialects and they can benefit from Chinese characters as alerting cues. The tonal monolinguals instead use a different orthography for English and do not benefit from Chinese characters as alerting cues.

Considering phonological activation, both phonological interference and phonological facilitation were found with the tonal bilinguals in Analysis 2 (also shown in Table 2), just as was found with the Chinese international students tested in previous studies (Li et al., 2013; Spinks et al., 2000). Thus, tonal bilinguals activate phonological information automatically in visual word recognition, just as the other Chinese speakers (Seidenberg, 1985; Tan & Perfetti, 1997; Zhou & Marslen-Wilson, 1999). However, the bilinguals did not show any evidence in support of tone-specific activation in any of the models. Neither were the effects from S+T+ characters significantly different from S+T− characters (Analysis 2 and Table 2), nor were the effects from S−T+ characters significantly different from S−T− characters (Analysis 3 and Table 3.1), which is different from the small tonal effect found on the tonal monolinguals (Analysis 4 and Table 3.2) and the previous monolingual findings in a picture-word interference experiment (a similar paradigm) (Nixon et al., 2014). Moreover, the between-dialect tonal relation showed no bilingual effects, with neither Analyses 5 nor 6 providing any support for cross-dialect tonal activation. It is unlikely that the tonal bilinguals do not make use of tonal information in lexical access, since we found plenty of within-dialect and between-dialect tonal effects on the same group of bilinguals (Wu & Chen, 2014; Wu et al., 2014, in prep). The lack of tonal activation may rather be related to the control of attention. For these tonal bilinguals, the two pronunciations of the same Chinese character are usually only different in tone. It could be that the tonal bilinguals, when reading silently and recognizing visual words automatically, maintain a shallower phonological processing compared with tonal monolinguals. Such arrangement can help the bilinguals avoid tone-based bilingual lexical competition and retrieve the lexical meaning more economically.

The tonal bilinguals also showed some scattered differences across the two language modes. The bilinguals' Stroop facilitation showed different patterns in different language modes (as shown in Analyses 1& 2 and Table 2). When the colors were named in SC, it was the color characters that triggered the greatest Stroop facilitation, similar to previous findings (Li et al., 2013; MacLeod, 1991; Spinks et al., 2000) and very reasonable, because the color characters share both phonological and semantic information with the congruent ink color. However, when the colors were named in JM, it was the S+T+ homophones that triggered the greatest Stroop facilitation, even greater than the Stroop facilitation triggered by the color characters. This finding is unusual and needs further replication. Is it possible that the semantic information of the Chinese characters was not activated in JM mode? This is unlikely because in that case the color characters would be equivalent to the S+T+ homophones and no difference in RT should emerge between the two conditions. Since the Stroop facilitation from the color characters still existed but was reduced, it is more likely that the additional semantic information from the color characters introduced interference in JM mode. This effect can be either interpreted semantically or in light of interlingual lexical competition. The color character represents exactly the same concept as the ink color (e.g. *RED* for the red ink). Many studies have shown that they are different from the semantic associates, which in Stroop-like tasks sometimes trigger semantic interference and sometimes do not (Finkbeiner & Caramazza, 2006; Starreveld & La Heij, 1995). If the reduced facilitation from the color characters is due to semantic interference, similar interference should also appear on the semantic associates. Since the current study did not include characters, which are semantic associates of color names as Spinks et al. (2000) did, further research is necessary to clarify this question. The interlingual lexical competition hypothesis is another explanation for the reduction of facilitation from the color characters. The Chinese character first activates the SC representation of the color name, which is different from the JM representation the tonal bilingual is trying to produce according to the language mode. The SC and JM translation equivalents, one activated by the color character and the other activated by the ink color, compete for lexical selection and counterweigh part of the phonological facilitation. The color character activates the unwanted SC representation of the color word via both the semantic route and the phonological route, but the S+T+ homophones of color characters only activate the same unwanted SC representation via the phonological route. Thus the unwanted SC representation of the color word is activated more by the color characters and introduced more interference in the JM color naming. This findings interesting, considering the fact that most tonal bilinguals received literacy education in SC and were not formally taught to read in JM. However, all of them can read aloud fluently in JM. How they manage to read Chinese characters in JM remains unclear but their automatic visual word recognition in the JM mode is obviously different from that in the SC mode. It could be that the visual word recognition in JM is actually mediated by the SC visual word recognition but not vice versa. This hypothesis can explain the asymmetrical facilitation from color characters in different language modes. When the bilinguals see a Chinese character, the SC representation is always automatically activated, but the JM representation needs mediation.

To sum up, the current study used the Stroop paradigm to investigate the role of lexical tones in automatic visual word recognition by a particular type of bilingual, namely the tonal bilinguals of two closely related Chinese dialects which use the same orthography and segmental structures for the translation equivalents. Beside the classical bilingual lexical disadvantage and executive advantage, our findings also corroborated the theory that phonological information carried by Chinese characters is automatically activated in visual word recognition, which is in line with previous findings. The tonal bilinguals showed similar sensitivity to segmental information as the tonal monolinguals. However, the tonal bilinguals seem to reduce their depth of visual language processing specifically in the tonal aspect, which would help them to a large extent avoid the conflicting tonal information from different dialects. Moreover, the dialect, which is rarely used in literacy education showed an unusual pattern in Stroop facilitation and this finding is discussed in light of bilingual lexical competition. Our study investigated previously neglected possibilities of bilingual visual word recognition and our findings provide new links between the process of lexical prosody and bilingual executive control.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11. doi: 10.1017/S1366728908003477

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859. doi: 10.1037/0278-7393.34.4.859

Biederman, I., & Tsao, Y.-C. (1979). On processing Chinese ideographs and English words: Some implications from Stroop-test results. *Cognitive Psychology, 11*(2), 125-132. doi: 10.1016/0010-0285(79)90007-0

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] Version 5.3.82 retrieved 26 July 2014 from http://www.praat.org/.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Chen, H.-C., & Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(3), 397. doi: 10.1037/0278-7393.12.3.397

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological review, 97*(3), 332. doi: 10.1037/0033-295X.97.3.332

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135-149. doi: j.cognition.2009.08.001

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59-86. doi: j.cognition.2006.12.013

Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics, 11*(4), 274-276. doi: 10.3758/BF03210377

Dennis, I., & Newstead, S. E. (1981). Is phonological recoding under strategic control? *Memory & Cognition, 9*(5), 472-477. doi: 10.3758/BF03202341

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior, 10*(3), 297-302. doi: 10.1016/S0022-5371(71)80057-9

Effler, M. (1978). Colour-distant versus colour-congruent colour words, their influence on naming times in the Stroop-test. *Psychologische Beitrage*.

Fang, S.-P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition, 9*(6), 609-617. doi: 10.3758/BF03202355

Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42*(6), 790-796. doi: 10.1016/S0010-9452(08)70419-2

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220-1234. doi: 10.3758/BF03193224

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review, 18*(4), 625-658. doi: 10.3758/s13423-011-0116-7

Kiyak, H. A. (1982). Interlingual interference in naming color words. *Journal of Cross-Cultural Psychology, 13*(1), 125-135. doi: 10.1177/0022022182131011

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://CRAN.R-project.org/package=lmerTest

Lee, T. M. C., & Chan, C. C. H. (2000). Stroop interference in Chinese and English. *Journal of Clinical and Experimental Neuropsychology, 22*(4), 465-471. doi: 10.1076/1380-3395(200008)22:4;1-0;FT465

Li, C., Lin, C., Wang, M., & Jiang, N. (2013). The activation of segmental and tonal information in visual word recognition. *Psychonomic bulletin & review*. doi: 10.3758/s13423-013-0395-2

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Intergrative Review. *Psychological Bulletin, 109*(2), 163-203. doi: 10.1037/0033-2909.109.2.163

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 126. doi: 10.1037/0278-7393.14.1.126

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., & Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain and Language, 120*(1), 61-65. doi: 10.1016/j.bandl.2011.10.003

Nixon, J. S., Chen, Y., & Schiller, N. O. (2014). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. *Language, Cognition and Neuroscience, 30*(5), 491-505. doi: 10.1080/23273798.2014.942326

Pacilly, J. J. A. (2010). *MarkIntervals.praat [Computer program] 24-jan-2010*.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633. doi: 10.1037/0278-7393.17.4.633

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24. doi: 10.1037/0278-7393.21.1.24

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium*: Lawrence Erlbaum.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 295-301. doi: 10.1016/S0022-5371(69)80079-4

Prior, A., & Gollan, T. H. (2011). Good Language-Switchers are Good Task-Switchers: Evidence from Spanish-English and Mandarin- English Bilinguals. *Journal of the International Neuropsychological Society, 17*(4), 682. doi: 10.1017/S1355617711000580

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26*(4), 392-405. doi: 10.1016/0749-596X(87)90098-2

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27*(3), 465. doi: 10.1017/S014271640606036X

Saalbach, H., & Stern, E. (2004). Differences between Chinese morphosyllabic and German alphabetic readers in the Stroop interference effect. *Psychonomic bulletin & review, 11*(4), 709-715. doi: 10.3758/BF03196624

SAS. (1978). Tests of Hypotheses in Fixed-Effects Linear Models *SAS Technical Report R-101* Cary, NC, USA: SAS Institute Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19*(1), 1-30. doi: 10.1016/0010-0277(85)90029-0

Singer, M. H., Lappin, J. S., & Moore, L. P. (1975). The interference of various word parts on color naming in the Stroop test. *Perception & Psychophysics, 18*(3), 191-193. doi: 10.3758/BF03205966

Smith, M. C., & Kirsner, K. (1982). Language and orthography as irrelevant features in colour-word and picture-word Stroop interference. *The Quarterly Journal of Experimental Psychology, 34*(1), 153-170. doi: 10.1080/14640748208400864

Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: The role of phonological information. *Cognition, 76*(1), B1-B11. doi: 10.1016/S0010-0277(00)00072-X

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 21(3), 686. doi: 10.1037/0278-7393.21.3.686

Stroop, J. R. (1935). Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology, 18*, 643-662. doi: 10.1037/h0054651

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*(1), 41-57. doi: 10.1006/jmla.1997.2508

Tsao, Y.-C., Wu, M.-F., & Feustel, T. (1981). Stroop interference: Hemispheric difference in Chinese speakers. *Brain and Language, 13*(2), 372-378. doi: 10.1016/0093-934X(81)90101-2

Underwood, G., & Briggs, P. (1984). The development of word recognition processes. *British Journal of Psychology, 75*(2), 243-255. doi: 10.1111/j.2044-8295.1984.tb01896.x

van der Loo, M. P. J. (2010). Distribution based outlier detection for univariate data,. *Discussion paper 10003, Statistics Netherlands, The Hague*.

Van Heuven, W. J. B., Conklin, K., Coderre, E. L., Guo, T., & Dijkstra, T. (2011). The influence of cross-language similarity on within-and between-language Stroop effects in trilinguals. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00374

Wu, J., & Chen, Y. (2014). *Tonal variants in the bilingual mental lexicon.* Paper presented at the The Fourth International Symposium on Tonal Aspects of Languages (TAL 2014), Nijmegen.

Wu, J., Chen, Y., Van Heuven, V. J., & Schiller, N. O. (2014). Tonal variability in lexical access. *Language, Cognition and Neuroscience*. doi: 10.1080/23273798.2014.915977

Wu, J., Chen, Y., Van Heuven, V. J., & Schiller, N. O. (in prep). Tonal similarity effect: the role of tone in the auditory lexical access of etymologically related translation equivalent.

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language, 41*, 579-606. doi: 10.1006/jmla.1999.2663

Appendices. Summary of the mixed effects models

Appendix 1. The model comparing the color characters with the neutral characters on LogRTs of congruent trials

Fixed effects	Df	F	p
Within-dialect condition (color characters/ neutral characters)	**1**	**14.658**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	3.099	<0.05
Color (blue/ green/ purple/red/yellow)	4	127.074	<0.001
Trial order in the same color (centralized)	1	27.212	<0.001
Trial distance from the same color (centralized)	1	39.222	<0.001
Within-dialect condition ×Mode	2	7.87	<0.001
Within-dialect condition ×Color	4	5.966	<0.001
Mode ×Color	8	2.765	<0.01
Within-dialect condition ×Trial order in the same color (centralized)	1	0.161	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	0.098	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	1.506	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	3.052	<0.05
Color ×Trial distance from the same color (centralized)	4	1.235	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.329	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	0.572	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	2.703	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	2.064	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized)	8	2.666	<0.01
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	0.238	>0.05(ns)
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	0.926	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	0.736	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	2.822	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.073	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	1.741	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	0.71	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.215	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.64	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	1.089	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.491	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.678	>0.05(ns)
Random effects		χ^2	
(1 + Trial distance from the same color (centralized) | Participant)	2	10.165	<0.01

Appendix 2. The model comparing the (S+T+) characters with the neutral characters on LogRTs of congruent trials

Fixed effects	Df	F	p
Within-dialect condition (homophones of the color characters/neutral characters	**1**	**32.339**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	3.567	<0.05
Color (blue/ green/ purple/red/yellow)	4	151.408	<0.001
Trial order in the same color (centralized)	1	18.061	<0.001
Trial distance from the same color (centralized)	1	67.744	<0.001
Within-dialect condition ×Mode	2	6.913	<0.01
Within-dialect condition ×Color	4	2.805	<0.05
Mode ×Color	8	2.043	<0.05
Within-dialect condition ×Trial order in the same color (centralized)	1	0.027	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	1.314	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	3.012	<0.05
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.624	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	1.385	>0.05(ns)
Color ×Trial distance from the same color (centralized)	4	1.177	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.09	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	1.352	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.417	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	0.375	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized)	8	2.415	<0.05
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	4.572	<0.05
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	1.141	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	0.919	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	4.246	<0.05
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.038	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.545	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.951	<0.05
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.761	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.555	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	1.453	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.984	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.678	>0.05(ns)
Random effects		χ^2	
(1 + Trial order in the same color (centralized) | Participant)	2	7.944	<0.05

Appendix 3. The model comparing the (S+T−) with the neutral characters on LogRTs of congruent trials

Fixed effects	Df	F	p
Within-dialect condition (different-tone homophones of the color characters/ neutral characters)	**1**	**21.133**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	6.336	<0.01
Color (blue/ green/ purple/red/yellow)	4	178.456	<0.001
Trial order in the same color (centralized)	1	15.071	<0.001
Trial distance from the same color (centralized)	1	40.07	<0.001
Within-dialect condition ×Mode	2	2.957	>0.05(ns)
Within-dialect condition ×Color	4	8.182	<0.001
Mode ×Color	8	1.321	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized)	1	0.988	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	2.255	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	0.945	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.013	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	0.541	>0.05(ns)
Color ×Trial distance from the same color (centralized)	4	0.876	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.579	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	1.995	<0.05
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.129	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	2.111	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized)	8	0.809	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	4.897	<0.01
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	0.554	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	1.468	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	2.244	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.304	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.282	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.844	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.462	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.272	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.275	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.267	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	1.575	>0.05(ns)
Random effects		χ^2	
(1 + Trial distance from the same color (centralized) | Participant)	2	11.690	<0.01

Appendix 4. The model comparing the JM tonal-merging characters with the neutral characters on LogRTs of congruent trials

Fixed effects	Df	F	p
Within-dialect condition (JM tonal-merging characters V.S. neutral characters)	**1**	**24.177**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	1.378	>0.05 (ns)
Color (blue/yellow)	1	188.615	<0.001
Trial order in the same color (centralized)	1	11.708	<0.001
Trial distance from the same color (centralized)	1	13.155	<0.001
Within-dialect condition ×Mode	2	0.766	>0.05 (ns)
Within-dialect condition ×Color	1	2.233	>0.05 (ns)
Mode ×Color	2	3.522	<0.05
Within-dialect condition ×Trial order in the same color (centralized)	1	2.905	>0.05 (ns)
Mode ×Trial order in the same color (centralized)	2	0.037	>0.05 (ns)
Color ×Trial order in the same color (centralized)	1	0.039	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.05	>0.05 (ns)
Mode ×Trial distance from the same color (centralized)	2	0.913	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	1	0.06	>0.05 (ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.453	>0.05 (ns)
Within-dialect condition ×Mode ×Color	2	0.13	>0.05 (ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.864	>0.05 (ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	1	3.552	>0.05 (ns)
Mode ×Color ×Trial order in the same color (centralized)	2	0.194	>0.05 (ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	4.488	<0.05
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	1	2.298	>0.05 (ns)
Mode ×Color ×Trial distance from the same color (centralized)	2	0.094	>0.05 (ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.079	>0.05 (ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.079	>0.05 (ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.166	>0.05 (ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	2	2.313	>0.05 (ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	2	1.968	>0.05 (ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.239	>0.05 (ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.002	>0.05 (ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.026	>0.05 (ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	2.747	>0.05 (ns)
Random effects		χ^2	
1 + Trial order in the same color (centralized) | Participant	2	14.210	<0.001

Appendix 5. The model comparing the color characters with the neutral characters on LogRTs of incongruent trials

Fixed effects	Df	F	p
Within-dialect condition (color characters/neutral characters)	**1**	**123.384**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	13.988	<0.001
Color (blue/ green/ purple/red/yellow)	4	32.009	<0.001
Trial order in the same color (centralized)	1	2.715	>0.05(ns)
Trial distance from the same color (centralized)	1	28.482	<0.001
Within-dialect condition ×Mode	2	21.591	<0.001
Within-dialect condition ×Color	4	1.969	>0.05(ns)
Mode ×Color	8	1.62	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized)	1	9.407	<0.01
Mode ×Trial order in the same color (centralized)	2	2.078	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	2.117	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.063	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	3.972	<0.05
Color ×Trial distance from the same color (centralized)	4	0.854	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.991	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	1.556	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	1.066	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	2.41	<0.05
Mode ×Color ×Trial order in the same color (centralized)	8	1.839	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	2.288	>0.05(ns)
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	0.315	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	1.292	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	1.545	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.606	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.642	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.657	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.798	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.061	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.671	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.262	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.819	>0.05(ns)
Random effects		χ^2	
1 + Trial order in the same color (centralized) | Participant	2	23.443	<0.001
1 + Trial distance from the same color (centralized) |StimuliID	2	10.389	<0.01

Appendix 6. The model comparing the (S+T+) characters with the neutral characters on LogRTs of incongruent trials

Fixed effects	Df	F	p
Within-dialect condition (homophones of the color characters /neutral characters)	**1**	**23.19**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	17.66	<0.001
Color (blue/ green/ purple/red/yellow)	4	56.548	<0.001
Trial order in the same color (centralized)	1	10.783	<0.01
Trial distance from the same color (centralized)	1	69.969	<0.001
Within-dialect condition ×Mode	2	6.038	<0.01
Within-dialect condition ×Color	4	2.015	>0.05(ns)
Mode ×Color	8	1.785	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized)	1	1.81	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	0.992	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	2.777	<0.05
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.08	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	2.775	>0.05(ns)
Color ×Trial distance from the same color (centralized)	4	2.636	<0.05
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.736	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	2.259	<0.05
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.672	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	2.401	<0.05
Mode ×Color ×Trial order in the same color (centralized)	8	1.63	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	4.453	<0.05
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	1.532	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	0.924	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	3.135	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	1.427	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.525	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.689	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	1.058	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.528	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.507	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.828	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	1.039	>0.05(ns)
Random effects		χ^2	
1 + Trial (centralized) | Participant	2	67.230	<0.001
1 | StimuliID	1	31.571	<0.001

Appendix 7. The model comparing the (S+T−) with the neutral characters on LogRTs of incongruent trials

Fixed effects	Df	F	p
Within-dialect condition (different-tone homophones of the color characters/ neutral characters)	**1**	**18.593**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	20.157	<0.001
Color (blue/ green/ purple/red/yellow)	4	52.573	<0.001
Trial order in the same color (centralized)	1	13.242	<0.001
Trial distance from the same color (centralized)	1	69.918	<0.001
Within-dialect condition ×Mode	2	5.929	<0.01
Within-dialect condition ×Color	4	0.728	>0.05(ns)
Mode ×Color	8	2.606	<0.01
Within-dialect condition ×Trial order in the same color (centralized)	1	3.74	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	2.237	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	2.452	<0.05
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.047	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	3.936	<0.05
Color ×Trial distance from the same color (centralized)	4	1.985	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	3.735	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	1.249	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.528	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	2.748	<0.05
Mode ×Color ×Trial order in the same color (centralized)	8	2.012	<0.05
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	1.963	>0.05(ns)
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	3.164	<0.05
Mode ×Color ×Trial distance from the same color (centralized)	8	1.239	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	1.77	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.983	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.375	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.369	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.47	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.591	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.458	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	1.04	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.432	>0.05(ns)
Random effects		χ^2	
1 + Trial (centralized) | Participant	2	77.775	<0.0001
1 | StimuliID	1	37.909	<0.001

Appendix 8. The model comparing the JM tonal-merging characters with the neutral characters on LogRTs of incongruent trials

Fixed effects	Df	F	p
Within-dialect condition (JM tonal-merging characters/neutral characters)	**1**	**37.024**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	5.335	<0.01
Color (blue/ green/ purple/red/yellow)	4	173.313	<0.001
Trial order in the same color (centralized)	1	7.635	<0.01
Trial distance from the same color (centralized)	1	39.554	<0.001
Within-dialect condition ×Mode	2	8.773	<0.001
Within-dialect condition ×Color	4	7.554	<0.001
Mode ×Color	8	3.863	<0.001
Within-dialect condition ×Trial order in the same color (centralized)	1	3.412	>0.05(ns)
Mode ×Trial order in the same color (centralized)	2	1.685	>0.05(ns)
Color ×Trial order in the same color (centralized)	4	0.529	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.803	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	3.048	<0.05
Color ×Trial distance from the same color (centralized)	4	1.432	>0.05(ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	2.994	>0.05(ns)
Within-dialect condition ×Mode ×Color	8	0.885	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized)	2	0.279	>0.05(ns)
Within-dialect condition ×Color ×Trial order in the same color (centralized)	4	3.939	<0.01
Mode ×Color ×Trial order in the same color (centralized)	8	1.446	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	5.046	<0.01
Within-dialect condition ×Color ×Trial distance from the same color (centralized)	4	1.802	>0.05(ns)
Mode ×Color ×Trial distance from the same color (centralized)	8	1.232	>0.05(ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	1.376	>0.05(ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.038	>0.05(ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.379	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized)	8	1.678	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial distance from the same color (centralized)	8	0.673	>0.05(ns)
Within-dialect condition ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	3.721	<0.05
Within-dialect condition ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	4	0.814	>0.05(ns)
Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.664	>0.05(ns)
Within-dialect condition ×Mode ×Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	8	0.666	>0.05(ns)
Random effects		χ^2	
1 + Trial order in the same color (centralized) | Participant	2	11.324	<0.01

Appendix 9. The model comparing different within-dialect conditions on LogRTs of congruent trials

Fixed effects	Df	F	p
Within-dialect condition color characters/ homophones of the color characters/ different-tone homophones of the color characters/ different-tone homophones of the color characters with potential tonal merging in JM)	**3**	**1.269**	**>0.05(ns)**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	9.502	<0.001
Trial (centralized)	1	10.868	<0.01
Trial distance from the same color (centralized)	1	88.171	<0.001
Within-dialect condition ×Mode	6	1.745	>0.05(ns)
Within-dialect condition ×Trial (centralized)	3	1.401	>0.05(ns)
Mode ×Trial (centralized)	2	0.363	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	3	0.708	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	2.532	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	3.714	>0.05(ns)
Within-dialect condition ×Mode ×Trial (centralized)	6	1.04	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	6	2.36	<0.05
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.697	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.18	>0.05(ns)
Within-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	6	0.404	>0.05(ns)
Random effects		χ^2	
1 +Trial (centralized) | Participant	2	44.346	<0.0001
1 | Color	1	1029.653	<0.001

Appendix 10. The model comparing different within-dialect conditions on LogRTs of incongruent trials

Fixed effects	Df	F	p
Within-dialect condition (color characters/ homophones of the color characters/ different-tone homophones of the color characters)	**3**	**79.161**	**<0.001**
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	58.136	<0.001
Trial (centralized)	1	3.123	>0.05(ns)
Trial distance from the same color (centralized)	1	227.712	<0.001
Within-dialect condition ×Mode	6	8.156	<0.001
Within-dialect condition ×Trial (centralized)	3	4.618	<0.01
Mode ×Trial (centralized)	2	2.543	>0.05(ns)
Within-dialect condition ×Trial distance from the same color (centralized)	3	1.104	>0.05(ns)
Mode ×Trial distance from the same color (centralized)	2	0.067	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	1.283	>0.05(ns)
Within-dialect condition ×Mode ×Trial (centralized)	6	0.632	>0.05(ns)
Within-dialect condition ×Mode ×Trial distance from the same color (centralized)	6	2.458	<0.05
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.532	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.864	>0.05(ns)
Within-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	6	1.512	>0.05(ns)
Random effects		χ^2	
1 +Trial (centralized) | Participant	2	189.357	<0.001
1 | StimuliID	1	163.154	<0.001
1 | Color	1	147.771	<0.001

Appendix 11. The model comparing the S−T− and S−T+ trials combined with blue ink color on LogRTs

Fixed effects	Df	F	p
Same Tone (the character is S−T− / S−T+ with the ink color)	2	2.081	>0.05 (ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	1	2.118	>0.05 (ns)
Trial (centralized)	1	0.11	>0.05 (ns)
Trial distance from the same color (centralized)	2	37.134	<0.001
Same Tone ×Mode	1	0.918	>0.05 (ns)
Same Tone ×Trial (centralized)	2	7.991	<0.01
Mode ×Trial (centralized)	1	1.687	>0.05 (ns)
Same Tone ×Trial distance from the same color (centralized)	2	0.121	>0.05 (ns)
Mode ×Trial distance from the same color (centralized)	1	1.314	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	2	0.101	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized)	2	0.656	>0.05 (ns)
Same Tone ×Mode ×Trial distance from the same color (centralized)	1	0.662	>0.05 (ns)
Same Tone ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.69	>0.05 (ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	1.093	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	1.548	>0.05 (ns)
Random effects		χ^2	
1+Trial (centralized) |Participant	2	1375	<0.01
1|StimuliID	1	5.89	<0.05

Appendix 12. The model comparing the S−T− and S−T+ trials combined with green ink color on LogRTs

Fixed effects	Df	F	p
Same Tone (the character is S−T− / S−T+ with the ink color)	1	0.053	>0.05 (ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	5.848	<0.01
Trial (centralized)	1	0.291	>0.05 (ns)
Trial distance from the same color (centralized)	1	34.151	<0.001
Same Tone ×Mode	2	2.757	>0.05 (ns)
Same Tone ×Trial (centralized)	1	0.562	>0.05 (ns)
Mode ×Trial (centralized)	2	0.21	>0.05 (ns)
Same Tone ×Trial distance from the same color (centralized)	1	1.21	>0.05 (ns)
Mode ×Trial distance from the same color (centralized)	2	4.423	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.01	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized)	2	0.518	>0.05 (ns)
Same Tone ×Mode ×Trial distance from the same color (centralized)	2	2.139	>0.05 (ns)
Same Tone ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.317	>0.05 (ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.799	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	3.901	<0.05
Random effects		χ^2	
1+ Trial distance from the same color (centralized)|Participant	2	6.49	<0.05
1|StimuliID	1	13.79	<0.0001

Appendix 13. The model comparing the S−T− and S−T+ trials combined with red ink color on LogRTs

Fixed effects	Df	F	p
Same Tone (the character is S−T− / S−T+ with the ink color)	1	2.1887	>0.05 (ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	2.3996	>0.05 (ns)
Trial (centralized)	1	0.3809	<0.05
Trial distance from the same color (centralized)	1	5.8	>0.05 (ns)
Same Tone ×Mode	2	2.0872	<0.05
Same Tone ×Trial (centralized)	1	5.6343	>0.05 (ns)
Mode ×Trial (centralized)	2	2.4177	>0.05 (ns)
Same Tone ×Trial distance from the same color (centralized)	1	0.607	>0.05 (ns)
Mode ×Trial distance from the same color (centralized)	2	0.9421	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.0464	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized)	2	1.7124	>0.05 (ns)
Same Tone ×Mode ×Trial distance from the same color (centralized)	2	0.6642	>0.05 (ns)
Same Tone ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.0092	>0.05 (ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.6429	>0.05 (ns)
Same Tone ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.3248	>0.05 (ns)
Random effects		χ^2	
1+ Trial (centralized)|Participant	2	2.60	0.273
1|StimuliID	1	3.71	0.054

Appendix 14. The model comparing the S−T− and S−T+ trials combined with yellow ink color on LogRTs

Fixed effects	Df	F	p
Same Tone (the character is S−T− / S−T+ with the ink color)	1	27.6397	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	4.449	<0.05
Trial order in the same color (centralized)	1	2.7757	>0.05 (ns)
Trial distance from the same color (centralized)	1	28.1628	<0.001
Same Tone ×Mode	2	1.0661	>0.05 (ns)
Same Tone ×Trial order in the same color (centralized)	1	2.7956	>0.05 (ns)
Mode ×Trial order in the same color (centralized)	2	4.0379	<0.05
Same Tone ×Trial distance from the same color (centralized)	1	1.3057	>0.05 (ns)
Mode ×Trial distance from the same color (centralized)	2	1.4558	>0.05 (ns)
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.0178	>0.05 (ns)
Same Tone ×Mode ×Trial order in the same color (centralized)	2	0.4635	>0.05 (ns)
Same Tone ×Mode ×Trial distance from the same color (centralized)	2	0.3022	>0.05 (ns)
Same Tone ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.0487	>0.05 (ns)
Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	0.4977	>0.05 (ns)
Same Tone ×Mode ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	2	1.6626	>0.05 (ns)
Random effects		χ^2	
1+ Trial order in the same color (centralized)|Participant	2	8.00	<0.05

Appendix 15. The model comparing the S+T− character 览 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in JM mode

Fixed effects	Df	F	p
Color (T+ color 'purple'/ T− color-'green'/ 'red'/ 'yellow')	3	34.29	<0.001
Within-dialect condition (S+T− character 览/ neutral character 鼻 for purple, 涂 for 'green', 贯 for 'red', 岸 for 'yellow')	1	0.148	>0.05 (ns)
Trial (centralized)	1	12.052	<0.001
Trial distance from the same color (centralized)	1	4.609	<0.05
Color ×Within-dialect condition	3	0.883	>0.05 (ns)
Color ×Trial (centralized)	3	1.248	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	0.031	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	1.703	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.657	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.046	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	0.774	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	2.491	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.11	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.782	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.771	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	138.18	<0.001

Appendix 16. The model compare the S+T− character 览 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in SC-bilingual mode

Fixed effects	Df	F	p
Color (T+ color 'purple'/ T− color × 'green'/ 'red'/ 'yellow')	3	67.205	<0.001
Within-dialect condition (S+T− character 览/ neutral character 鼻 for purple, 涂 for 'green', 贯 for 'red', 岸 for 'yellow')	1	0.061	>0.05 (ns)
Trial (centralized)	1	2.35	>0.05 (ns)
Trial distance from the same color (centralized)	1	14.697	<0.001
Color ×Within-dialect condition	3	1.834	>0.05 (ns)
Color ×Trial (centralized)	3	1.088	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	3.09	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	0.473	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	4.184	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	2.166	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	0.156	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	0.554	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.109	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	1.283	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.751	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	101.25	<0.001

Appendix 17. The model comparing the S+T− character 览 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in SC-monolingual mode

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘green’/ ‘red’/ ‘yellow’)	3	45.057	<0.001
Within-dialect condition (S+T− character 览/ neutral character 鼻 for purple, 涂 for ‘green’, 贯 for ‘red’, 岸 for ‘yellow’)	1	13.535	<0.001
Trial (centralized)	1	0.083	>0.05 (ns)
Trial distance from the same color (centralized)	1	2.919	>0.05 (ns)
Color ×Within-dialect condition	3	2.099	>0.05 (ns)
Color ×Trial (centralized)	3	1.441	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	0.67	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	0.071	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.12	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.75	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	5.007	<0.01
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	1.222	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.534	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	1.717	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	1.331	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	-1.81	1

Appendix 18. The model comparing the S+T− character 旅 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent JM trials

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘blue’/ ‘red’/ ‘yellow’)	3	51.874	<0.001
Within-dialect condition (S+T− character 旅/ neutral character 鼻 for purple, 抱 for ‘blue’, 贯 for ‘red’, 岸 for ‘yellow’)	1	3.333	>0.05 (ns)
Trial (centralized)	1	7.403	<0.01
Trial distance from the same color (centralized)	1	10.452	<0.01
Color ×Within-dialect condition	3	1.398	>0.05 (ns)
Color ×Trial (centralized)	3	3.146	<0.05
Within-dialect condition ×Trial (centralized)	1	2.28	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	2.896	<0.05
Within-dialect condition ×Trial distance from the same color (centralized)	1	4.508	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.041	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	1.893	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	1.384	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	3.089	<0.05
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.304	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.904	>0.05 (ns)
Random effects		χ^2	
1+ Trial (centralized)|Participant	2	13.39	<0.01

Appendix 19. The model comparing the S+T− character 旅 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent SC-bilingual trials

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘blue’/ ‘red’/ ‘yellow’)	3	70.994	<0.001
Within-dialect condition (S+T− character 旅/ neutral character 鼻 for purple, 抱 for ‘blue’, 贯 for ‘red’, 岸 for ‘yellow’)	1	1.739	>0.05 (ns)
Trial (centralized)	1	5.737	<0.05
Trial distance from the same color (centralized)	1	20.489	<0.001
Color ×Within-dialect condition	3	0.64	>0.05 (ns)
Color ×Trial (centralized)	3	1.261	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	1.721	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	1.431	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.551	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.223	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	0.155	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	0.429	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.428	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	3.89	<0.05
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	1.408	>0.05 (ns)
Random effects		χ^2	
1+ Trial (centralized)|Participant	2	3.85	>0.05

Appendix 20. The model comparing the S+T− character 旅 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in SC-monolingual mode

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘blue’/ ‘red’/ ‘yellow’)	3	67.551	<0.001
Within-dialect condition (S+T− character 旅/ neutral character 鼻 for purple, 抱 for ‘blue’, 贯 for ‘red’, 岸 for ‘yellow’)	1	8.297	<0.01
Trial order in the same color (centralized)	1	1.615	>0.05 (ns)
Trial distance from the same color (centralized)	1	20.563	<0.001
Color ×Within-dialect condition	3	1.092	>0.05 (ns)
Color ×Trial order in the same color (centralized)	3	0.78	>0.05 (ns)
Within-dialect condition ×Trial order in the same color (centralized)	1	0.245	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	0.226	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	10.876	<0.01
Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	0.006	>0.05 (ns)
Color ×Within-dialect condition ×Trial order in the same color (centralized)	3	6.803	<0.001
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	0.062	>0.05 (ns)
Color ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	3	0.281	>0.05 (ns)
Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	1	2.289	>0.05 (ns)
Color ×Within-dialect condition ×Trial order in the same color (centralized) ×Trial distance from the same color (centralized)	3	1.528	>0.05 (ns)
Random effects		χ^2	
1+ Trial order in the same color (centralized) | Participant	1	13.34	<0.001

Appendix 21. The model comparing the S+T− character 谎 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in JM mode

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘blue’/’green’/ ‘red’/ ‘yellow’)	3	36.214	<0.001
Within-dialect condition (S+T− character 谎/ neutral character 抱 for ‘blue’, 涂 for ‘green’, 鼻 for ‘purple’, 贯 for ‘red’)	1	12.642	<0.001
Trial (centralized)	1	3.596	>0.05 (ns)
Trial distance from the same color (centralized)	1	11.386	<0.001
Color ×Within-dialect condition	3	1.858	>0.05 (ns)
Color ×Trial (centralized)	3	1.022	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	0.826	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	3.872	<0.01
Within-dialect condition ×Trial distance from the same color (centralized)	1	1.681	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.468	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	1.448	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	2.459	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.139	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	1.012	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.845	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	83.66	<0.001

Appendix 22. The model comparing the S+T− character 谎 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in SC-bilingual mode

Fixed effects	Df	F	p
Color (T+ color ‘purple’/ T− color × ‘blue’/’green’/ ‘red’/ ‘yellow’)	3	46.391	**<0.001**
Within-dialect condition (S+T− character 谎/ neutral character 抱 for ‘blue’, 涂 for ‘green’, 鼻 for ‘purple’, 贯 for ‘red’)	1	9.389	**<0.01**
Trial (centralized)	1	7.158	<0.01
Trial distance from the same color (centralized)	1	12.665	<0.001
Color ×Within-dialect condition	3	1.664	>0.05 (ns)
Color ×Trial (centralized)	3	1.062	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	1.897	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	0.906	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	5.441	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.014	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	0.351	>0.05 (ns)
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	0.221	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.409	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	2.355	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.642	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	126.16	<0.001

Appendix 23. The model compare the S+T− character 谎 combined with T− versus T+ incongruent ink colors on LogRTs of incongruent trials in SC-monolingual mode

Fixed effects	Df	F	p
Color (T+ color 'purple'/ T− color × 'blue'/'green'/ 'red'/ 'yellow')	3	70.72	<0.001
Within-dialect condition (S+T− character 谎/ neutral character 抱 for 'blue', 涂 for 'green', 鼻 for 'purple', 贯 for 'red')	1	54.164	<0.001
Trial (centralized)	1	0.056	>0.05 (ns)
Trial distance from the same color (centralized)	1	0.845	>0.05 (ns)
Color ×Within-dialect condition	3	4.503	<0.01
Color ×Trial (centralized)	3	0.666	>0.05 (ns)
Within-dialect condition ×Trial (centralized)	1	0.479	>0.05 (ns)
Color ×Trial distance from the same color (centralized)	3	0.753	>0.05 (ns)
Within-dialect condition ×Trial distance from the same color (centralized)	1	0.647	>0.05 (ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.003	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized)	3	2.683	<0.05
Color ×Within-dialect condition ×Trial distance from the same color (centralized)	3	0.595	>0.05 (ns)
Color ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.49	>0.05 (ns)
Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	2.825	>0.05 (ns)
Color ×Within-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	3	0.111	>0.05 (ns)
Random effects		χ^2	
1 | Participant	1	9.68	<0.01

Appendix 24. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of incongruent trials with blue ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(烂)/ S+T− false friend of the corresponding color character (览))	1	39.628	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	3.571	<0.05
Trial (centralized)	1	4.392	<0.05
Trial distance from the same color (centralized)	1	20.986	<0.001
Between-dialect condition ×Mode	**2**	0.084	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	1	0.009	>0.05(ns)
Mode×Trial (centralized)	2	1.281	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	0.039	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	0.867	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.557	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	1.17	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	1.91	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.326	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	1.551	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.166	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	816.78	<0.001
1 | Color	1	354.59	<0.001

Appendix 25. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of congruent trials with blue ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(烂)/ S+T− false friend of the corresponding color character (览))	1	2.6773	>0.05(ns)
Mode	2	0.2365	>0.05(ns)
Trial (centralized)	1	2.743	>0.05(ns)
Trial distance from the same color (centralized)	1	4.1986	<0.05
Between-dialect condition ×Mode	2	0.8759	>0.05(ns)
Between-dialect condition ×Trial (centralized)	1	0.4364	>0.05(ns)
Mode×Trial (centralized)	2	1.7094	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	0	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	1.2275	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.004	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	0.8356	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	1.8123	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	2.0186	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.1947	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.0801	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	15.15	<0.001

Appendix 26. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of incongruent trials with yellow ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(荒)/ S+T− false friend of the corresponding color character (谎))	1	5.5308	<0.01
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	3.6248	<0.01
Trial (centralized)	1	1.7352	>0.05(ns)
Trial distance from the same color (centralized)	1	20.6441	<0.001
Between-dialect condition ×Mode	2	0.7057	>0.05(ns)
Between-dialect condition ×Trial (centralized)	1	0.1695	>0.05(ns)
Mode×Trial (centralized)	2	2.2653	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	0.0134	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	2.915	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.1867	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	1.6649	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	6.4513	<0.01
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	2.363	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.1425	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	5.1322	<0.01
Random effects		χ^2	
1 | Participant	1	876.06	<0.001
1 | Color	1	309.16	<0.001

Appendix 27. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of congruent trials with yellow ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(荒)/ S+T− false friend of the corresponding color character (谎))	1	2.6096	>0.05(ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	1.1472	>0.05(ns)
Trial (centralized)	1	0.021	>0.05(ns)
Trial distance from the same color (centralized)	1	13.638	<0.001
Between-dialect condition ×Mode	2	3.1512	<0.05
Between-dialect condition ×Trial (centralized)	1	1.4148	>0.05(ns)
Mode×Trial (centralized)	2	1.4	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	0.1018	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	1.4791	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.0046	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	0.7688	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	0.4793	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.2155	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.6039	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.7844	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	12.010	<0.001

Appendix 28. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of incongruent trials with green ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(驴)/ between language homophones with different tone (旅))	1	33.462	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	3.598	<0.05
Trial (centralized)	1	4.017	<0.05
Trial distance from the same color (centralized)	1	51.676	<0.001
Between-dialect condition ×Mode	**2**	**1.866**	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	1	0.137	>0.05(ns)
Mode×Trial (centralized)	2	0.204	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	2.189	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	4.527	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	1.514	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	0.324	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	0.84	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.008	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.349	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	1.587	>0.05(ns)
Random effects		χ^2	
1 + Trial (centralized)| Participant	2	7.193	<0.05
1 | Color	1	473.81	<0.001

Appendix 29. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of congruent trials with green ink color

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character(驴)/ between language homophones with different tone (旅))	1	0.3051	>0.05(ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	1.0845	>0.05(ns)
Trial (centralized)	1	2.598	>0.05(ns)
Trial distance from the same color (centralized)	1	8.0267	<0.01
Between-dialect condition ×Mode	**2**	**0.8536**	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	1	1.3967	>0.05(ns)
Mode×Trial (centralized)	2	2.8052	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	1	4.0011	<0.05
Mode×Trial distance from the same color (centralized)	2	1.885	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.8296	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	2	2.4631	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	2	2.6056	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	1	0.5199	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	2.6901	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.0838	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	-1.21	1

Appendix 30. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of incongruent color character trials

Fixed effects	Df	F	p
Between-dialect condition (S+T+ color cognates ‘green’/ S+T+ color cognates ×’ blue/ purple/ red/ yellow’)	4	1.512	>0.05(ns)
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	9.876	<0.001
Trial (centralized)	1	1.099	>0.05(ns)
Trial distance from the same color (centralized)	1	57.701	<0.001
Between-dialect condition ×Mode	8	1.401	>0.05(ns)
Between-dialect condition ×Trial (centralized)	4	1.519	>0.05(ns)
Mode×Trial (centralized)	2	2.288	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	4	3.071	<0.05
Mode×Trial distance from the same color (centralized)	2	2.808	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	1.27	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	8	1.832	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	8	0.762	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	4	0.159	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.647	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	8	0.721	>0.05(ns)
Random effects		χ^2	
1+Trial.crl|Participant	2	7.389	<0.05
1|StimuliID	1	15.197	<0.001
1 | Color	1	32.916	<0.001

Appendix 31. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of congruent color character trials

Fixed effects	Df	F	p
Between-dialect condition (S+T+ color cognates ‘green’/ S+T+ color cognates ×’ blue/ purple/ red/ yellow’)	4	58.867	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	1.108	>0.05(ns)
Trial (centralized)	1	13.332	<0.001
Trial distance from the same color (centralized)	1	23.378	<0.001
Between-dialect condition ×Mode	**8**	**1.056**	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	4	1.281	>0.05(ns)
Mode×Trial (centralized)	2	1.369	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	4	0.953	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	0.964	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	0.758	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	8	1.143	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	8	0.365	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	4	2.23	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.141	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	8	0.433	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	15.937	<0.001

Appendix 32. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of incongruent S+T+ trials

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character 虑/ S+T− false friend of the corresponding color character ×栏/ 子/ 洪/ 皇)	4	15.594	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	13.914	<0.001
Trial (centralized)	1	6.389	<0.05
Trial distance from the same color (centralized)	1	100.192	<0.001
Between-dialect condition ×Mode	**8**	**1.452**	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	4	2.148	>0.05(ns)
Mode×Trial (centralized)	2	1.168	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	4	0.181	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	0.491	>0.05(ns)
Trial (centralized) ×Trial distance from the same color (centralized)	1	7.228	<0.01
Between-dialect condition ×Mode ×Trial (centralized)	8	0.774	>0.05(ns)
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	8	1.682	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	4	5.288	<0.001
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	1.165	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	8	1.63	>0.05(ns)
Random effects		χ^2	
1+Trial.crl|Participant	2	38.49	<0.001
1 | Color	1	1173.1	<0.001

Appendix 33. The model investigating the interaction of Mode and the between-dialect condition on LogRTs of congruent S+T+ trials

Fixed effects	Df	F	p
Between-dialect condition (S+T+ false friend of the corresponding color character 虑/ S+T− false friend of the corresponding color character ×栏/子/洪/ 皇)	4	75.856	<0.001
Mode (SC by Beijing monolinguals/ JM/ SC by Jinan bilinguals)	2	1.49	>0.05(ns)
Trial (centralized)	1	8.233	<0.01
Trial distance from the same color (centralized)	1	43.733	<0.001
Between-dialect condition ×Mode	**8**	**0.829**	**>0.05(ns)**
Between-dialect condition ×Trial (centralized)	4	1.217	>0.05(ns)
Mode×Trial (centralized)	2	0.4	>0.05(ns)
Between-dialect condition ×Trial distance from the same color (centralized)	4	0.14	>0.05(ns)
Mode×Trial distance from the same color (centralized)	2	3.792	<0.05
Trial (centralized) ×Trial distance from the same color (centralized)	1	3.593	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized)	8	2.905	<0.01
Between-dialect condition ×Mode ×Trial distance from the same color (centralized)	8	1.135	>0.05(ns)
Between-dialect condition ×Trial (centralized) ×Trial distance from the same color (centralized)	4	1.61	>0.05(ns)
Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	2	0.001	>0.05(ns)
Between-dialect condition ×Mode ×Trial (centralized) ×Trial distance from the same color (centralized)	8	1.146	>0.05(ns)
Random effects		χ^2	
1 | Participant	1	80.488	<0.001

8 General Discussion

This thesis investigated the role of tone in lexical processing by tonal bilinguals of two related Chinese dialects, Standard Chinese (SC) and Jinan Mandarin (JM). Theories and findings on different aspects of lexical processing were revisited in the context of this special type of bilingualism. In the following sections, the findings of individual chapters are first summarized, and then their general implications are discussed.

8.1 Summary for individual chapters

Chapter 2 investigates to what extent the interlingual category-goodness between different tonal systems keeps its impact on lexical access and speech comprehension. The acoustic distributions, interlingual perception, and the corresponding interlingual lexical and semantic activations of SC and JM rising tones were investigated in a group of experiments. An asymmetry was found in production and perception: the JM rising tone is more similar to the SC high-rising tone in acoustic distribution and interlingual perception. Such asymmetry also influences the tonal bilinguals' interlingual lexical access: SC high-rising final pseudo-words were more likely to be accepted as a JM real word than their low-rising final counterparts. However, the asymmetrical mapping did not affect the semantic activation after lexical access, which suggests some discreteness in the different levels of speech comprehension, namely between lexical activation and semantic activation. Interlingual two-to-one mapping is known to affect speech perception (Best & Strange, 1992; Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997; Miyawaki et al., 1975) and lexical access (Cutler & Otake, 2004; Dufour, Nguyen, & Frauenfelder, 2007; Pallier, Colomé, & Sebastián-Gallés, 2001). Chapter 2 generalized the previous findings to the tonal system of native tonal bilinguals and expanded the research questions to include lexical access and the corresponding word comprehension.

Chapter 3 addresses the question of how the strength of systematic correspondence is influenced by the sociolinguistic and cognitive backgrounds of the individuals. Between-word pitch distances of JM words can be predicted from SC tonal categories using statistical modeling, with interlingual tonal identity and individual backgrounds taken into consideration. The global influence of the bilinguals' sociolinguistic and cognitive backgrounds on the strength of systematic correspondence was found for the first time. The expected success of this statistical prediction mainly verified what historical linguists have known for decades, namely the systematic correspondence mechanism (Dyen, 1963; Meillet & Ford, 1967). Moreover, this study revealed the way sociolinguistic and cognitive backgrounds affect the general strength of systematic correspondence for the first time. The age-dependent and -independent effects were statistically teased apart.

What is the role of tonal similarity in the auditory lexical access of etymologically related translation equivalents? Chapter 4 tapped into this question and found that, although the bilinguals showed no difference in the general reaction times to tonal-identical and tonal-non-identical translation equivalents, tonal similarity showed different effects (facilitatory vs. interfering) on the lexical access of tonal-identical and tonal-non-identical translation equivalents. The interaction of tonal similarity with language mode (i.e. SC versus JM) and test blocks was also complex. Regardless of the status of the translation equivalents (i.e. identical or not), the SC word was processed faster than its JM equivalent with SC as the dominant dialect. This indicates that the lexical representations of translation equivalents in the bilingual lexicon are not only distinguished by lexical tones but also by language mode. Moreover, the SC-JM bilinguals showed an unusual lexical advantage compared to the tonal-monolingual controls, suggesting that bilinguals of closely related dialects, with their mental lexicon full of etymologically related translation equivalents, may benefit from this structure of their bilingual mental lexicon in auditory lexical access. The present finding is different from earlier findings using visual tasks (Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010), where the discontinuous effects of tonal similarity in visual word recognition suggested that the orthographically identical cognates share one common lexical representation. Complex interactions were found between tonal similarity and language mode. This suggests that language mode influences how exactly tonal similarity affects the lexical co-activation and lexical competition between etymologically related translation equivalents.

Chapter 5 studied how tonal pattern variation between tonal lexical variants affects auditory lexical access. True repetition, within-category variation, tonal pattern variation, and lexically contrastive variation were compared in an auditory form-priming experiment. The results support the view that tonal patterns have representative status in lexical access but also converge in a lexically specific way. Different types of variability have been investigated for decades in models of auditory lexical recognition and speech production (Connine, Ranbom, & Patterson, 2008; Goldinger, 1998; Lahiri & Reetz, 2002). The present study contributes to this topic by providing new evidence from the suprasegmental aspect. By including a previously untested condition these findings are in line with earlier studies of pronunciation variants, and support that both tonal-pattern variants are stored under the same lemma (Bürki, Ernestus, & Frauenfelder, 2010; Ernestus, 2014) and are sensitive to variant frequency (Connine et al., 2008). However, the present findings generally cannot be explained by the underspecification hypothesis of lexical representation, or listeners' tolerance of mismatches in the process (Lahiri & Marslen-Wilson, 1991; Marslen-Wilson & Zwitserlood, 1989). This is because a general flexibility regarding tone would predict a lack of difference across different types of tonal conditions, which counters the findings in Chapter 5.

Chapter 6 compared JM lexical variants which are either identical or non-identical to their SC translation equivalents and examined their lexical access for speech production. The variant probability effect suggests that the SC-JM bilinguals do store the JM tonal lexical variant they did not produce in the corpus, and that the difference in individual one-time choice still mainly reflects the variant probability instead of the individual difference in lexical representation. The variant probability effect was found in tonal variants for the first time and the effect is in line with the segmental findings (Connine et al., 2008).

Do SC-JM tonal bilinguals differ from SC tonal monolinguals when automatically retrieving tonal information from Chinese characters? What tonal information does the common written form of the etymologically related translation equivalents activate? In order to answer these questions, Chapter 7 adopted a Stroop paradigm. Although showing a general lexical disadvantage in visual word production, the bilinguals benefited more from the congruent conditions and suffered less from the incongruent conditions in this Stroop experiment. The SC-JM bilinguals also differ from the SC tonal monolinguals in their lack of tonal sensitivity. Although phonological interference existed in both the tonal monolinguals and the tonal bilinguals, only the tonal monolinguals showed tonal effects. The present findings are different from the results of previous Chinese Stroop experiments (Li, Lin, Wang, & Jiang, 2013; Spinks, Liu, Perfetti, & Tan, 2000), which did not specify whether their Chinese participants had tonal backgrounds other than Standard Chinese and found incongruent tonal effects. Our findings suggest that tonal bilinguals are different from tonal monolinguals in their tone-related attention-control. Future studies on Chinese visual word recognition should take the participants' experiences with tonal dialects into consideration.

8.2 General implications

8.2.1 The role of tone in lexical process

Lexical tones distinguish words in tonal languages. Just like lexical stress (Levelt, Roelofs, & Meyer, 1999; Schiller & Costa, 2006), lexical tones function like abstract lexical frames in lexical access. However, not all the tone-related aspects of lexical process are clear. Additionally, findings from tonal languages may potentially inspire the studies on lexical stress in return.

For instance, how are tonal lexical variants [i.e. /tɕiɛn tan (High-level+Rising)/ and /tɕiɛn tan (Low+High)/ both mean 'simple' in JM] stored and processed in lexical access? Chapter 5 adopted a form priming paradigm and investigated the role of different levels of tonal variability in a single-dialect auditory lexical decision experiment. In light of the findings from Chapter 5, Chapter 6 investigated the issue of tonal representation in the context of bilingualism. These chapters show that unproduced lexical variants are nevertheless stored and the naming latency is affected by the variant probability. Tonal patterns may have representative status, but the lexically specific convergence of tonal patterns also happens in lexical access. It would be interesting to see whether these findings also apply to other types of tonal languages with register tones (e.g. Yoruba) and pitch accent languages (e.g. Swedish). These findings may also be of interest for studies on languages with lexical variants which are different in stress [e.g. válisi and valísi both mean 'tooth' in the Budai Rukai dialect of Formosan (C.-M. Chen, 2006)].

8.2.2 Interlingual matching of phonological inventories

Tonal bilinguals face a series of dilemmas regarding the phonological processing before lexical retrieval, some of which can also exist between remote languages. For instance, one tonal category in Dialect A is similar to two different tonal categories in Dialect B. Can both of these tonal categories from Dialect B activate the same tonal category in Dialect A during lexical access? What is the effect of interlingual category-goodness? Does the effect of interlingual category-goodness last until semantic activation? Chapter 2 looked specifically into a tonal case of phonological similarity in the sound inventories. It is shown that, in auditory lexical access, the tonal acoustic space can be divided differently according to the language mode, which is sensitive to interlingual category-goodness, and the effect of interlingual category-goodness only lasts until lexical access occurs. These findings may be generalized to tonal bilingualism involving remote languages and to similar two-to-one matching cases involving consonants and vowels.

8.2.3 Mental representation of etymologically related translation equivalents

Some findings in this thesis are specific to etymologically related translation equivalents, including both cognates and loan-words. Previous studies mostly use 'cognate' to refer to all such words and the most well-known findings are about the 'cognate facilitation' effect (Costa, Caramazza, & Sebastian-Galles, 2000; Dijkstra, Grainger, & Van Heuven, 1999). Closely related dialects are special in that the bilingual mental lexicon is teaming with etymologically related translation equivalents. Bilingualism involving closely related dialects offers an ideal test case for the mental representation of etymologically related translation equivalents.

Chapter 3 shows that the systematic correspondence between SC-JM etymologically related words varies across bilingual individuals, affected by their sociolinguistic and cognitive backgrounds. Using an auditory lexical decision task, Chapter 6 shows that, in the bilingual mental lexicon, lexical nodes are distinguished not only by the pronunciation but also by the language mode, whether the translation equivalents are identical or not. This finding is not totally consistent with earlier findings (Dijkstra et al., 2010).

The inconsistency may come from several different sources, because the SC-JM case is even more special than just two closely related tonal dialects. First, there is segmental identity between the etymologically related translation equivalents. Second, the same written form (Chinese character) is used for the translation equivalents, which means that the previously used visual word recognition paradigms cannot be tested in the current case. In light of these unique features, what is the cause behind the inconsistent findings? It could be that the auditory and the visual route yield access to different underlying lexical representations. It could also be that a bilingual lexicon dominated by etymologically related translation equivalents functions differently from a bilingual lexicon where etymologically related translation equivalents only exist sporadically. Alternatively, it could be that tonal information is processed differently from segmental information in bilingual lexical access as well.

To answer the theoretical questions on the bilingual lexical representation of etymologically related words, many empirical questions still need to be addressed. If bilinguals can be found between two dialects using different written forms for the etymologically related translation equivalents (e.g. Urdu and Hindi), would the pattern of effect be more similar to previous findings or to the current findings? Would a bilingual lexicon with a mediocre number of etymologically related translation equivalents show a different pattern (e.g. SC with a more remote Chinese dialect)? If another pair of dialects is tested, whose etymologically related translation equivalents are tonally identical but vary in segmental similarity, would similar effects as found in the current study still show up? To further understand the lexical representation of etymologically related words, more bilingual cases with closely related dialects need to be tested.

8.2.4 Bilingual visual word recognition of logographic written forms

How phonological information is activated and retrieved via logographic written forms (especially Chinese characters) has intrigued researchers for many years (Perfetti & Zhang, 1991, 1995; Tzeng, Hung, & Wang, 1977; Wu, Zhou, & Shu, 1999; Zhou & Marslen-Wilson, 1999). Bilingual visual word recognition involving various types of writing systems has also received a lot of attention (H.-C. Chen & Ho, 1986; Dyer, 1971; Fang, Tzeng, & Alva, 1981; Kiyak, 1982; Preston & Lambert, 1969). It is surprising that few studies have taken into consideration that the same logographic written forms can be associated with different pronunciations in related dialects. It was assumed that findings from experiments, without knowing the participants' dialect backgrounds, could be generalized to both monolinguals and bilinguals who use this logographic writing system.

When reading Chinese characters aloud, SC-JM bilinguals' performance is influenced by variant-probability (Chapter 6) and individual backgrounds (Chapter 3). When simply coming across Chinese characters, these bilinguals also activate the relevant mental representations. Chapter 7 shows that the participants' dialectal backgrounds affect their automatic visual word recognition of the common logographic written forms. The pattern of Stroop effects differs between SC-JM tonal bilinguals and SC tonal monolinguals. Though neither group showed exactly the same pattern as found in earlier Chinese Stroop experiments, it was the Stroop facilitation and interference found in the tonal bilingual group, surprisingly, that aligned better with earlier findings (Li et al., 2013; Spinks et al., 2000).

Note that in the current case of bilingualism, the same written form is associated with the same segmental structure but potentially different tonal contours. If the same written form is instead associated with similar but different segmental structures (e.g. for a bilingual who speak both SC and Shanghai Wu dialect), how would the automatic phonological activation happen in automatic visual word recognition? For instance, would both pronunciations be automatically activated? If so, would the activation happen simultaneously or sequentially? Further studies are necessary in order to answer these questions.

8.2.5 Tonal bilinguals' advantages and disadvantages

The bilingual advantages and disadvantages related to lexical tasks were also revisited in this thesis. Bilinguals are known to have lexical disadvantages and executive-control advantages. On the one hand, most previous studies showed that bilinguals are slower in lexical access compared with monolinguals (Bialystok, 2009; Martin et al., 2012; Ransdell & Fischler, 1987). This lexical disadvantage was explained by the fact that bilinguals have a denser lexical neighborhood and hence suffer from more lateral inhibitions (Ransdell & Fischler, 1987). On the other hand, bilinguals are better at resolving conflicts in tasks (Bialystok, 2009; Carlson & Meltzoff, 2008; Hilchey & Klein, 2011; Prior & Gollan, 2011), including the conflict in the Stroop task (Bialystok, Craik, & Luk, 2008). This executive advantage can be explained by the fact that bilinguals receive more training in language switching, which is a conflict-resolution process. Earlier evidence for bilingual advantages and disadvantages has mostly been found in bilinguals of non-tonal languages. Are these earlier findings applicable to tonal bilinguals of closely related dialects?

Regarding bilingual executive-control advantages, the answer to this question is ‘yes’. Tonal bilinguals benefited more from the congruent conditions and suffered less from the incongruent conditions (Chapter 7) compared with the monolinguals in the Stroop experiment. This result is consistent with earlier findings (Bialystok et al., 2008). Moreover, the new Stroop findings also draw special attention to bilingual attention control, especially regarding tone. Taking into consideration the findings of Chapters 5 & 6, it is apparent that tone is important for the tonal bilinguals in distinguishing lexical contrasts and lexical variants in auditory and production tasks. Tonal information is aurally available in auditory lexical recognition, and necessary for reading aloud. In contrast, when the bilinguals need to name the ink colors instead of the word itself, the tonal information of the written word becomes distracting. Thus, the lack of tonal sensitivity may actually be of advantage to the tonal bilinguals in this Stroop paradigm. Indeed, the tonal bilinguals are able to redirect their attention away from tone in the automatic visual word activation, better than tonal monolinguals. Recently, the theory of bilingual executive-control advantages was challenged (Paap, 2015; Paap, Johnson, & Sawi, 2015). Paap (2015) said that these advantages are probably ‘restricted to very specific and undetermined circumstances’. The current case of tonal bilingualism indeed revealed a very specific executive advantage related to tone. However, the current study also clarifies the circumstance for this advantage.

Then, as for the bilingual lexical disadvantages, the answer to the question above is both ‘yes’ and ‘no’. Not only lexical disadvantage but also lexical advantage was found in this study. In automatic visual word recognition (Chapter 7), as expected, bilinguals named colors generally slower than monolinguals. However, in the auditory lexical decision of etymologically related translation equivalents (Chapter 4), the SC-JM bilinguals showed an unexpected advantage over SC tonal monolinguals in SC mode, after switching from a block in JM mode. One may argue that these bilinguals' lexical advantage may be only restricted to etymologically related translation equivalents. However, as bilinguals of two related tonal dialects, they have mental lexicons teaming with such words. In other words, this is their normality, not exception. Also, note that the SC-JM bilinguals and SC tonal monolinguals both know some English. Thus, the difference cannot be attributed to the knowledge of non-tonal foreign language. Both the lexical disadvantage and lexical advantage here are specific to the tonal bilingualism of related dialects.

More attention is needed for this type of bilingual lexicon, which is teaming with etymologically related translation equivalents. In earlier studies, even with the same cognates, the bilinguals were still mostly found to be slower than the monolinguals in cognate production (Costa et al., 2000) and visual word recognition (Dijkstra et al., 1999; Lemhöfer, Dijkstra, & Michel, 2004; Lemhöfer et al., 2008; Mulder, Dijkstra, Schreuder, & Baayen, 2014). As the JM-SC translation equivalents are always stored as separate lexical nodes (as found in Chapter 4), there is no reason to believe that the lexical neighborhood is less dense for the JM-SC tonal bilinguals. Then it needs to be considered that these bilinguals may have acquired special advantages in handling the etymologically related translation equivalents in their mental lexicon. For instance, the SC-JM tonal bilinguals may benefit more from lexical coactivation and suffer less from the lateral inhibition for such lexical nodes. This interpretation is also consistent with Chapter 7's findings on speech production, in that the SC-JM bilinguals showed greater Stroop facilitation in consistent conditions and smaller Stroop interference in inconsistent conditions than the SC monolinguals.

To conclude, the behavioral findings in this thesis provided empirical evidence on speech perception, lexical processing, and attention control of tonal bilinguals of related tonal dialects. It started with how tonal bilinguals process the two-to-one interlingual tonal mapping in bilingual lexical access and semantic activation. Then it moved on to how they handle different levels of tonal variability, how they make use of the tonal information to access the target lexical variant, and how they store tonally identical and non-identical translation equivalents. Furthermore, it investigated how these bilinguals benefit from the co-activation and fight with the competition of etymologically related translation equivalents, and use their fine-tuned attention and executive control to cope with task-irrelevant tonal information. Related theoretical issues are revisited and the new findings suggest that current models of bilingual lexical processing need to be adjusted to accommodate the possibilities provided by the tonal bilingualism of related dialects.

References

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11. doi: 10.1017/S1366728908003477

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 34(4), 859. doi: 10.1037/0278-7393.34.4.859

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(03), 303-328. doi: 10.1017/S0142716400008912

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one 'fenêtre' in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language, 62*(4), 421-437. doi: 10.1016/j.jml.2010.01.002

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Chen, C.-M. (2006) *A comparative study on Formosan phonology: Paiwan and Budai Rukai* (pp. 273): ProQuest.

Chen, H.-C., & Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(3), 397. doi: 10.1037/0278-7393.12.3.397

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics, 70*(3), 403-411.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283. doi: 10.1037/0278-7393.26.5.1283

Cutler, A., & Otake, T. (2004). Pseudo-homophony in non-native listening. *The Journal of the Acoustical Society of America, 115*(5), 2392-2392. doi: 10.1121/1.4780547

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301. doi: 10.1016/j.jml.2009.12.003

Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America, 121*(4). doi: 10.1121/1.2710742

Dyen, I. (1963). Why phonetic change is regular. *Language, 39*(4), 631-637. doi: 10.2307/411958

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior, 10*(3), 297-302. doi: 10.1016/S0022-5371(71)80057-9

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*(1), 27-41. doi: 10.1016/j.lingua.2012.12.006

Fang, S.-P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition, 9*(6), 609-617. doi: 10.3758/BF03202355

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470. doi: 10.1006/jpho.1997.0052

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review, 105*(2), 251.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review, 18*(4), 625-658. doi: 10.3758/s13423-011-0116-7

Kiyak, H. A. (1982). Interlingual interference in naming color words. *Journal of Cross-Cultural Psychology, 13*(1), 125-135. doi: 10.1177/0022022182131011

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245-294. doi: 10.1016/0010-0277(91)90008-R

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. *Laboratory phonology, 7*, 637-676.

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*(5), 585-611. doi: 10.1080/01690960444000007

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: a megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 12. doi: 10.1037/0278-7393.34.1.12

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behaviroral and Brain Sciences, 22*(1), 1-75. doi: 10.1017/S0140525X99001776

Li, C., Lin, C., Wang, M., & Jiang, N. (2013). The activation of segmental and tonal information in visual word recognition. *Psychonomic bulletin & review*. doi: 10.3758/s13423-013-0395-2

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 576-585. doi: 10.1037/0096-1523.15.3.576

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., & Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain and Language, 120*(1), 61-65. doi: 10.1016/j.bandl.2011.10.003

Meillet, A., & Ford, G. B. (1967). *The comparative method in historical linguistics*. Paris: H. Champion.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*(5), 331-340. doi: 10.1111/j.1551-6709.2010.01140.x

Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, H. R. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language, 72*, 59-84. doi: 10.1016/j.jml.2013.12.004

Paap, K. R. (2015). Do many hones dull the bilingual whetstone? *Bilingualism: Language and Cognition, 18*(01), 41-42. doi: 10.1017/S1366728914000431

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*(0). doi: http://dx.doi.org/10.1016/j.cortex.2015.04.014

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449. doi: 10.1111/1467-9280.00383

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633. doi: 10.1037/0278-7393.17.4.633

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24. doi: 10.1037/0278-7393.21.1.24

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 295-301. doi: 10.1016/S0022-5371(69)80079-4

Prior, A., & Gollan, T. H. (2011). Good Language-Switchers are Good Task-Switchers: Evidence from Spanish-English and Mandarin- English Bilinguals. *Journal of the International Neuropsychological Society, 17*(4), 682. doi: 10.1017/S1355617711000580

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26*(4), 392-405. doi: 10.1016/0749-596X(87)90098-2

Schiller, N. O., & Costa, A. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *The Mental Lexicon, 1*(2), 231-250. doi: 10.1075/ml.1.2.04sch

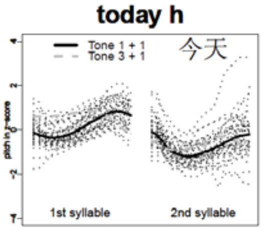
Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: The role of phonological information. *Cognition, 76*(1), B1-B11. doi: 10.1016/S0010-0277(00)00072-X

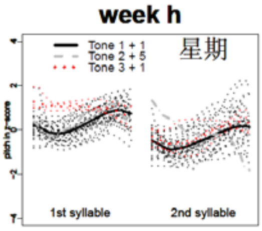
Tzeng, O. J., Hung, D. L., & Wang, W. S. Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory, 3*(6), 621-630. doi: 10.1037/0278-7393.3.6.621

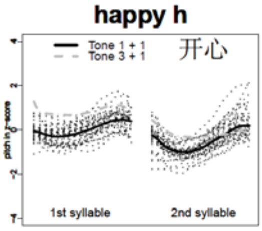
Wu, N., Zhou, X., & Shu, H. (1999). Sublexical processing in reading Chinese: A development study. *Language and Cognitive Processes, 14*(5), 503-524. doi: 10.1080/016909699386176

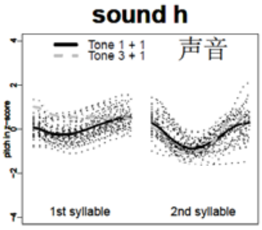
Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language, 41*, 579-606. doi: 10.1006/jmla.1999.2663

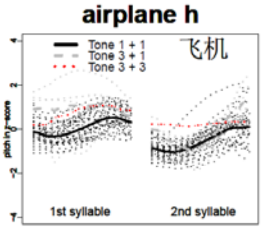
Appendices. Word-Wise Contour Plots

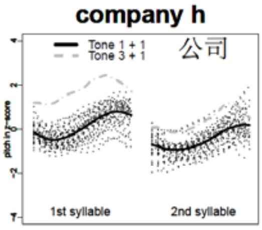


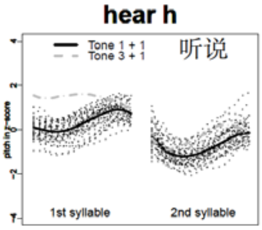


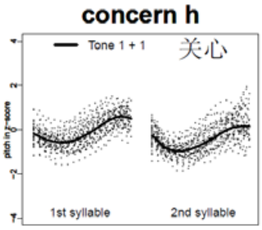


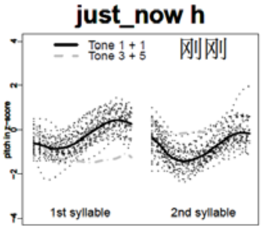


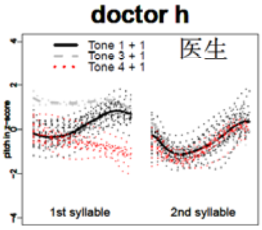


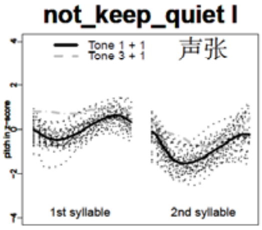


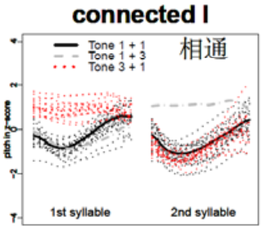


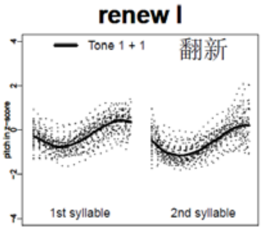


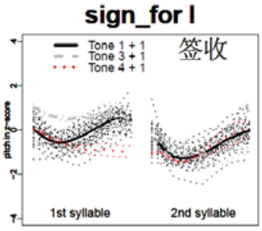


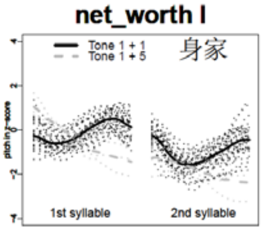


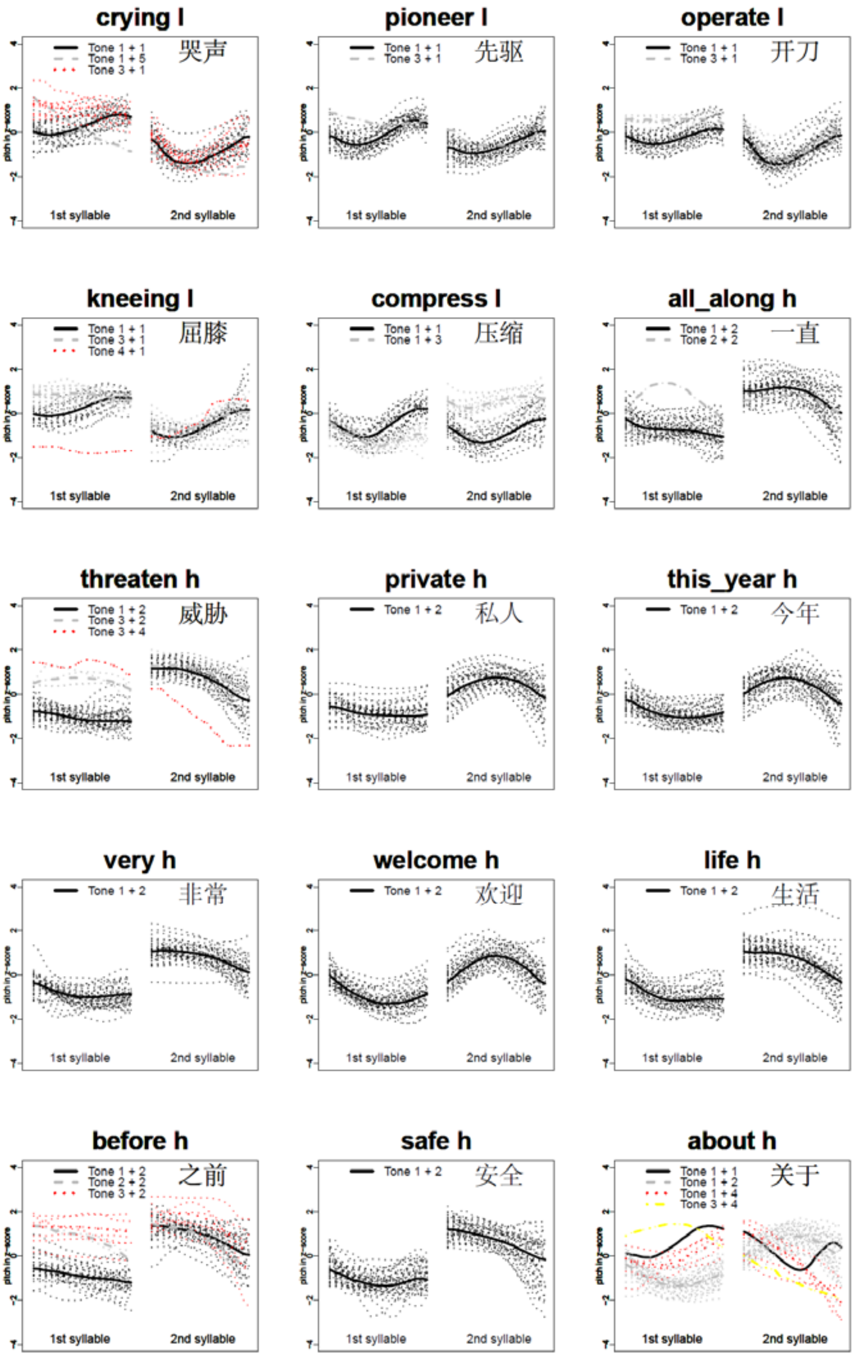


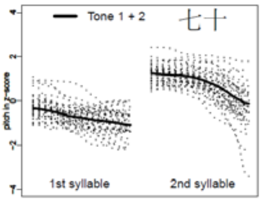


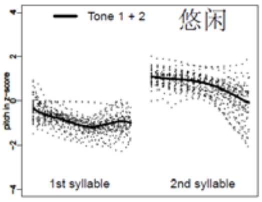


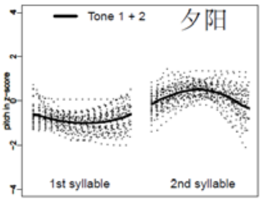


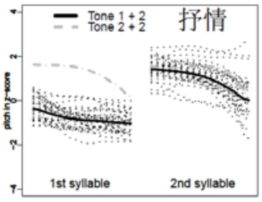


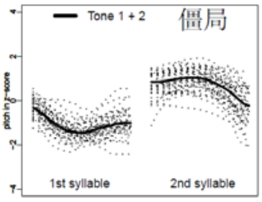

crying l
Tone 1 + 1
Tone 1 + 5
Tone 3 + 1
哭声
pioneer l
Tone 1 + 1
Tone 3 + 1
先驱
operate l
Tone 1 + 1
Tone 3 + 1
开刀
kneeing l
Tone 1 + 1
Tone 3 + 1
Tone 4 + 1
屈膝
compress l
Tone 1 + 1
Tone 1 + 3
压缩
all_along h
Tone 1 + 2
Tone 2 + 2
一直
threaten h
Tone 1 + 2
Tone 3 + 2
Tone 3 + 4
威胁
private h
Tone 1 + 2
私人
this_year h
Tone 1 + 2
今年
very h
Tone 1 + 2
非常
welcome h
Tone 1 + 2
欢迎
life h
Tone 1 + 2
生活
before h
Tone 1 + 2
Tone 2 + 2
Tone 3 + 2
之前
safe h
Tone 1 + 2
安全
about h
Tone 1 + 1
Tone 1 + 2
Tone 1 + 4
Tone 3 + 4
关于
pitch in z-score
-4
-2
0
2
4
1st syllable
2nd syllable

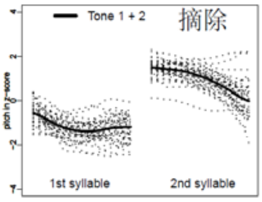

seventy l

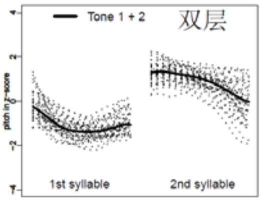

leisure l

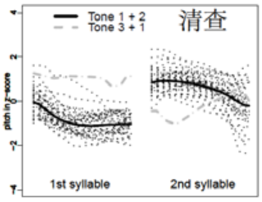

sunset l

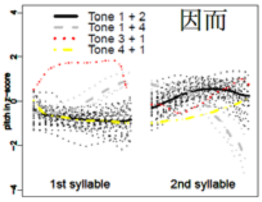

lyrical l

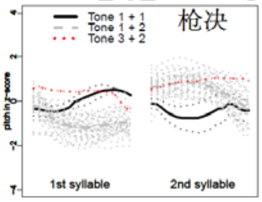

deadlock l

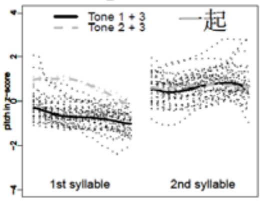

excise l

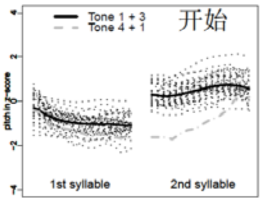

double l

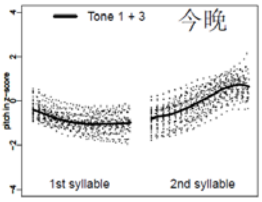

check l

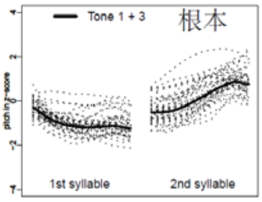

thus l

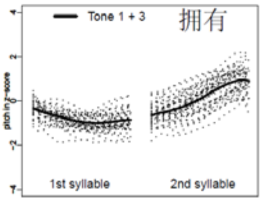

execute_by_shooting l

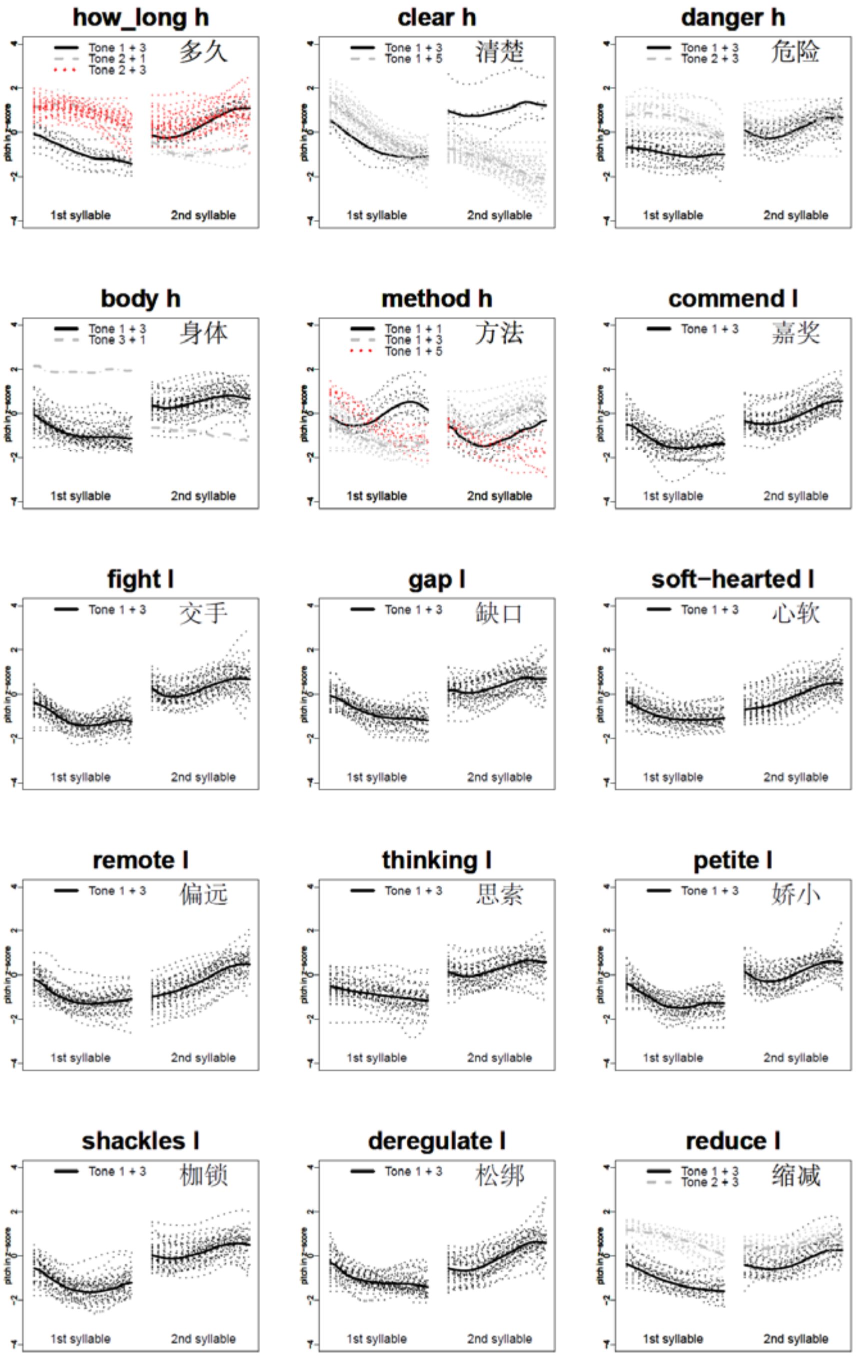

together h

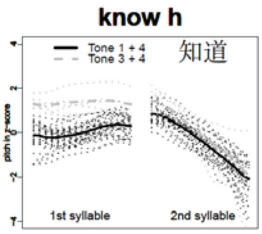

start h

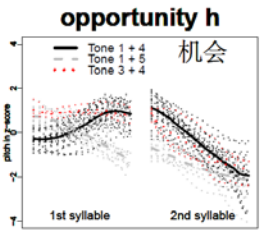

tonight h

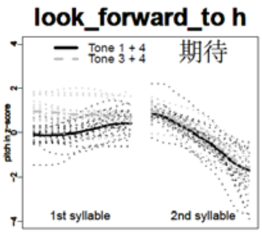

fundamental h

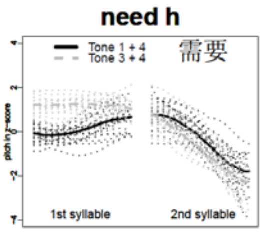

own h

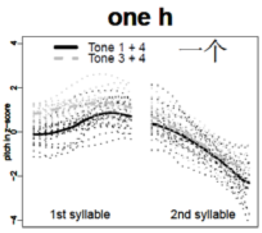

how_long h
Tone 1 + 3
Tone 2 + 1
Tone 2 + 3
多久
clear h
Tone 1 + 3
Tone 1 + 5
清楚
danger h
Tone 1 + 3
Tone 2 + 3
危险
body h
Tone 1 + 3
Tone 3 + 1
身体
method h
Tone 1 + 1
Tone 1 + 3
Tone 1 + 5
方法
commend l
Tone 1 + 3
嘉奖
fight l
Tone 1 + 3
交手
gap l
Tone 1 + 3
缺口
soft−hearted l
Tone 1 + 3
心软
remote l
Tone 1 + 3
偏远
thinking l
Tone 1 + 3
思索
petite l
Tone 1 + 3
娇小
shackles l
Tone 1 + 3
枷锁
deregulate l
Tone 1 + 3
松绑
reduce l
Tone 1 + 3
Tone 2 + 3
缩减
pitch in z−score
4
2
0
−2
−4
1st syllable
2nd syllable

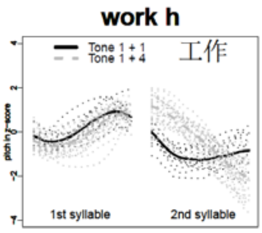

know h
Tone 1 + 4
Tone 3 + 4
知道
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

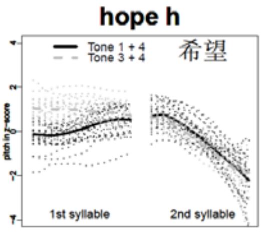

opportunity h
Tone 1 + 4
Tone 1 + 5
Tone 3 + 4
机会
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

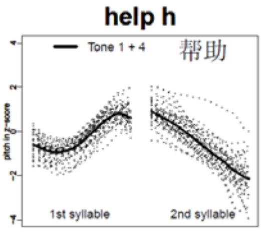

look_forward_to h
Tone 1 + 4
Tone 3 + 4
期待
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

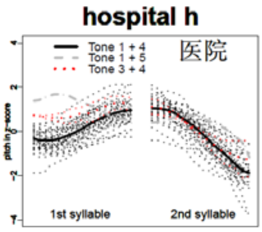

need h
Tone 1 + 4
Tone 3 + 4
需要
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

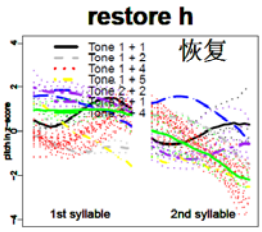

one h
Tone 1 + 4
Tone 3 + 4
一个
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

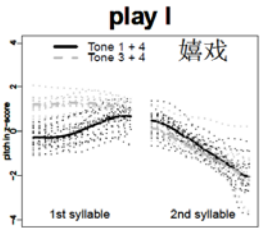

work h
Tone 1 + 1
Tone 1 + 4
工作
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

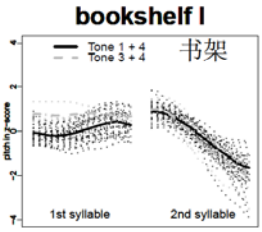

hope h
Tone 1 + 4
Tone 3 + 4
希望
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

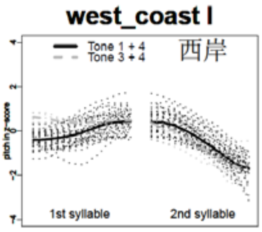

help h
Tone 1 + 4
帮助
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

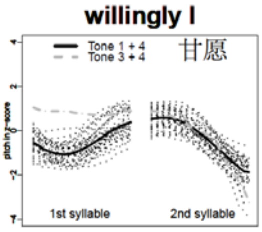

hospital h
Tone 1 + 4
Tone 1 + 5
Tone 3 + 4
医院
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

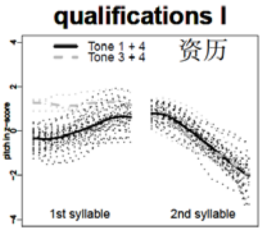

restore h
Tone 1 + 1
Tone 1 + 2
Tone 1 + 4
Tone 1 + 5
Tone 2 + 2
Tone 3 + 1
Tone 3 + 4
恢复
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

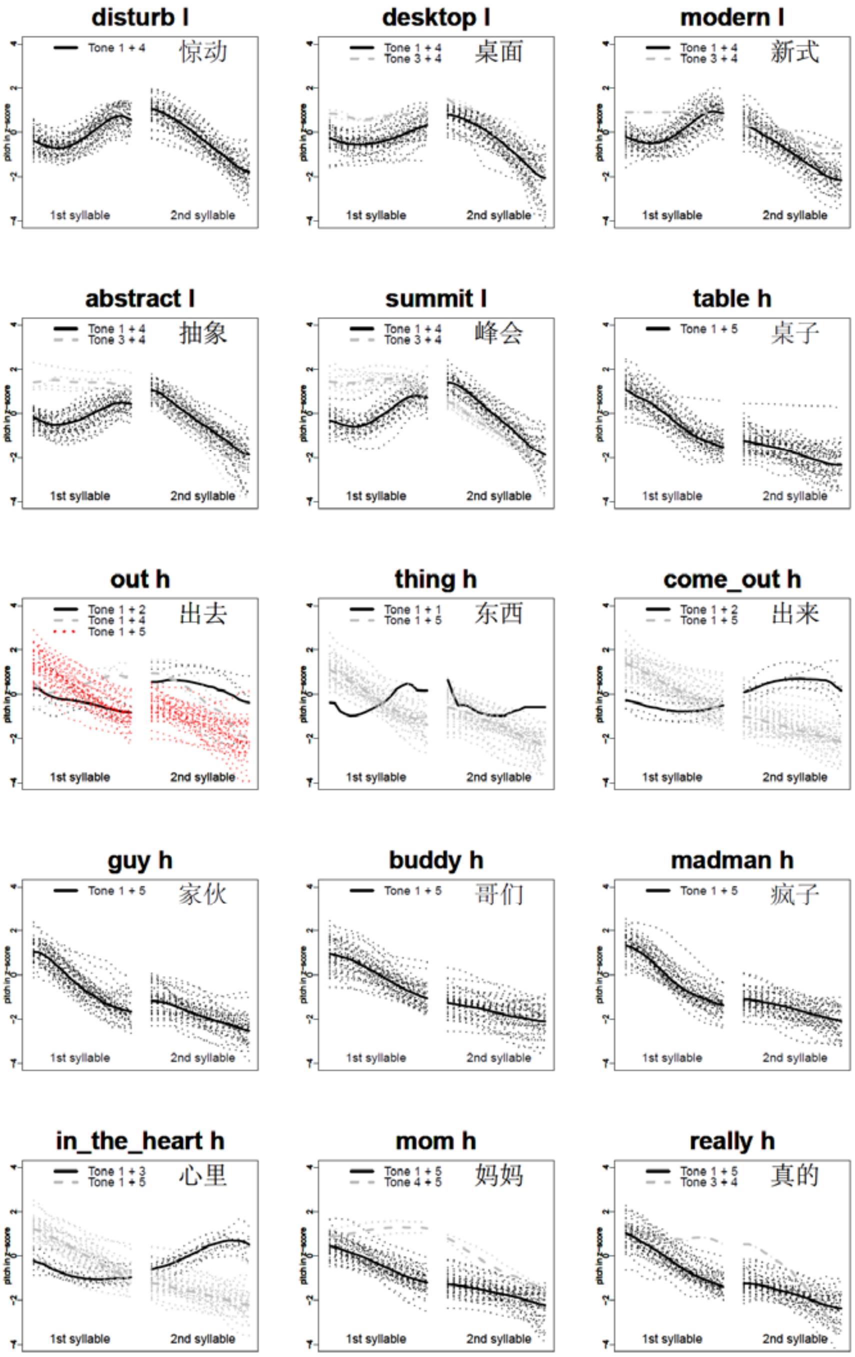

play l
Tone 1 + 4
Tone 3 + 4
嬉戏
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

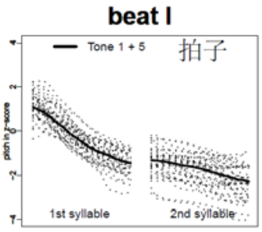

bookshelf l
Tone 1 + 4
Tone 3 + 4
书架
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

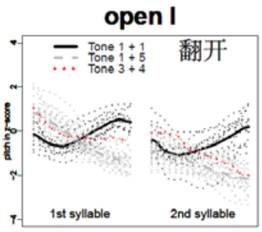

west_coast l
Tone 1 + 4
Tone 3 + 4
西岸
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

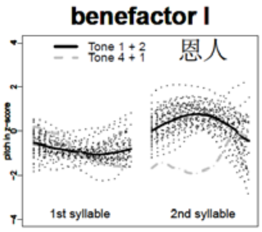

willingly l
Tone 1 + 4
Tone 3 + 4
甘愿
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

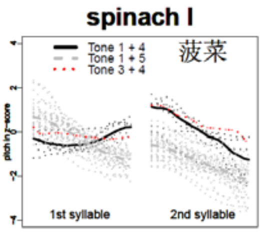

qualifications l
Tone 1 + 4
Tone 3 + 4
资历
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

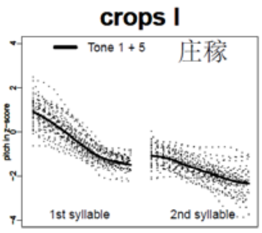

disturb l
Tone 1 + 4
惊动
desktop l
Tone 1 + 4
Tone 3 + 4
桌面
modern l
Tone 1 + 4
Tone 3 + 4
新式
abstract l
Tone 1 + 4
Tone 3 + 4
抽象
summit l
Tone 1 + 4
Tone 3 + 4
峰会
table h
Tone 1 + 5
桌子
out h
Tone 1 + 2
Tone 1 + 4
Tone 1 + 5
出去
thing h
Tone 1 + 1
Tone 1 + 5
东西
come_out h
Tone 1 + 2
Tone 1 + 5
出来
guy h
Tone 1 + 5
家伙
buddy h
Tone 1 + 5
哥们
madman h
Tone 1 + 5
疯子
in_the_heart h
Tone 1 + 3
Tone 1 + 5
心里
mom h
Tone 1 + 5
Tone 4 + 5
妈妈
really h
Tone 1 + 5
Tone 3 + 4
真的
pitch in z-score
1st syllable
2nd syllable

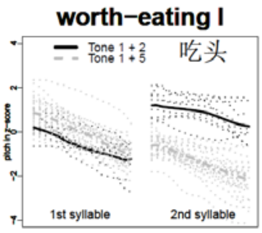

beat l
Tone 1 + 5
拍子
pitch in z-score
1st syllable
2nd syllable

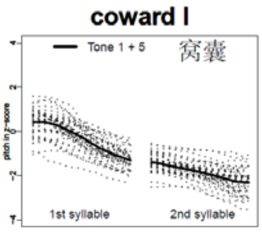

open l
Tone 1 + 1
Tone 1 + 5
Tone 3 + 4
翻开
pitch in z-score
1st syllable
2nd syllable

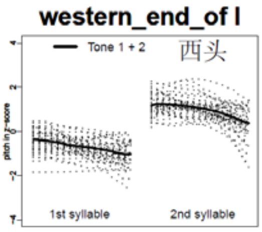

benefactor l
Tone 1 + 2
Tone 4 + 1
恩人
pitch in z-score
1st syllable
2nd syllable

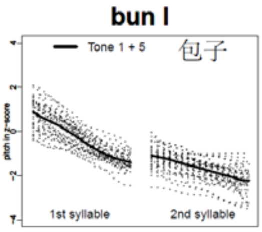

spinach l
Tone 1 + 4
Tone 1 + 5
Tone 3 + 4
菠菜
pitch in z-score
1st syllable
2nd syllable

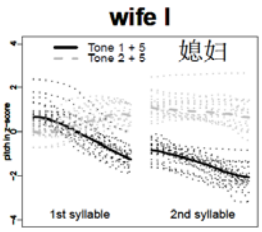

crops l
Tone 1 + 5
庄稼
pitch in z-score
1st syllable
2nd syllable

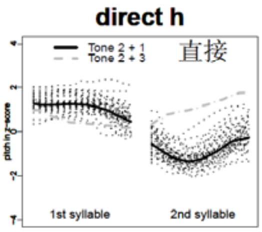

worth−eating l
Tone 1 + 2
Tone 1 + 5
吃头
pitch in z-score
1st syllable
2nd syllable

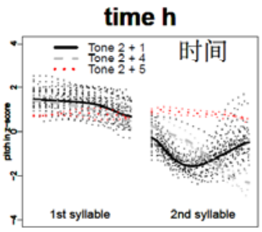

coward l
Tone 1 + 5
窝囊
pitch in z-score
1st syllable
2nd syllable

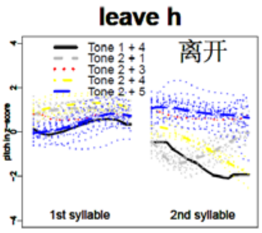

western_end_of l
Tone 1 + 2
西头
pitch in z-score
1st syllable
2nd syllable

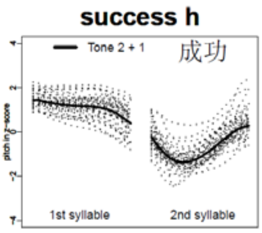

bun l
Tone 1 + 5
包子
pitch in z-score
1st syllable
2nd syllable

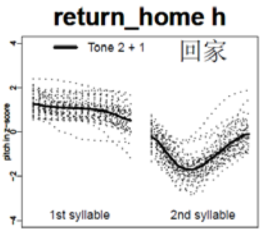

wife l
Tone 1 + 5
Tone 2 + 5
媳妇
pitch in z-score
1st syllable
2nd syllable

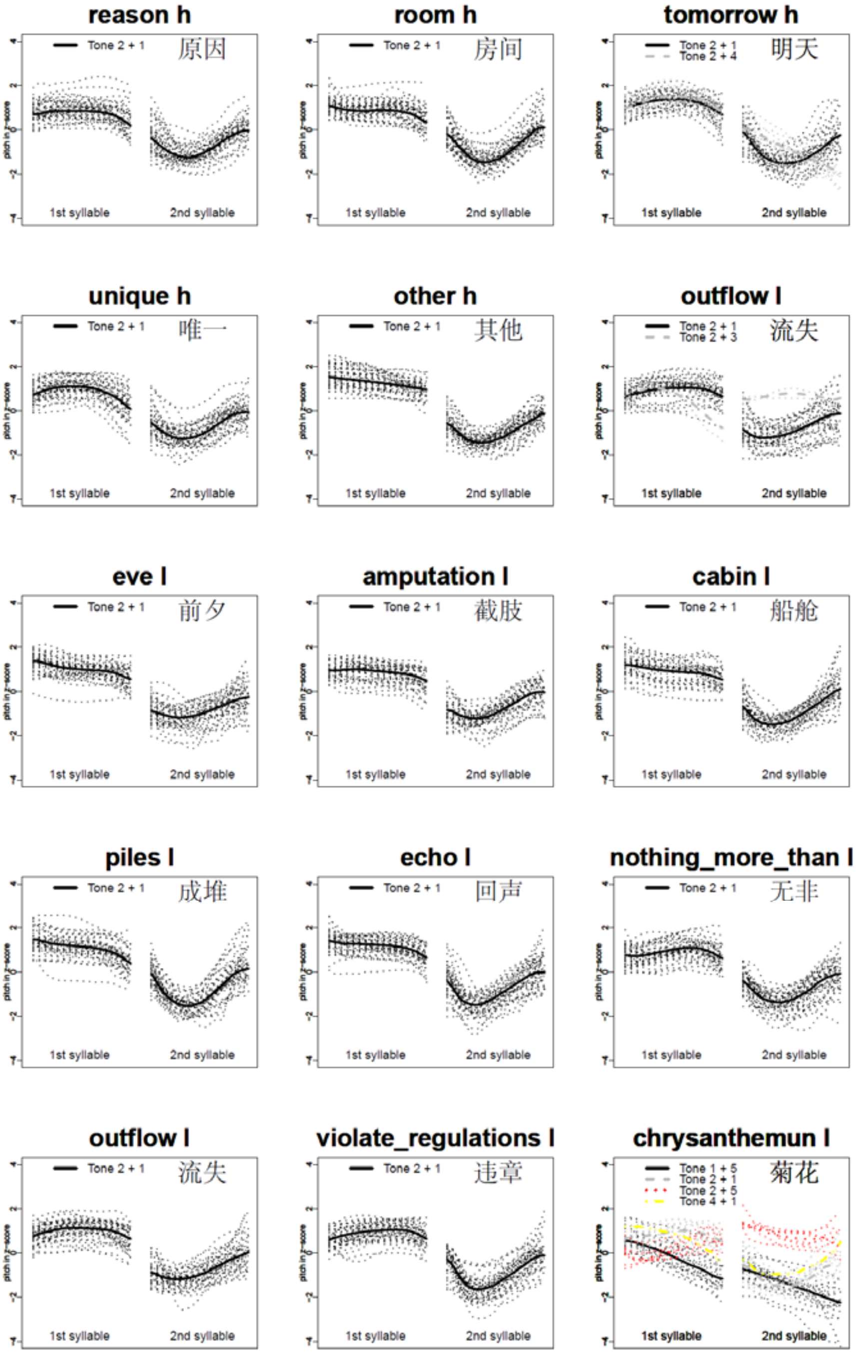

direct h
Tone 2 + 1
Tone 2 + 3
直接
pitch in z-score
1st syllable
2nd syllable


time h
Tone 2 + 1
Tone 2 + 4
Tone 2 + 5
时间
pitch in z-score
1st syllable
2nd syllable

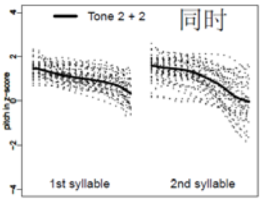

leave h
Tone 1 + 4
Tone 2 + 1
Tone 2 + 3
Tone 2 + 4
Tone 2 + 5
离开
pitch in z-score
1st syllable
2nd syllable


success h
Tone 2 + 1
成功
pitch in z-score
1st syllable
2nd syllable

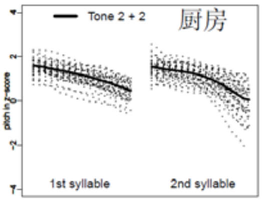

return_home h
Tone 2 + 1
回家
pitch in z-score
1st syllable
2nd syllable


reason h
Tone 2 + 1
原因
room h
Tone 2 + 1
房间
tomorrow h
Tone 2 + 1
Tone 2 + 4
明天
unique h
Tone 2 + 1
唯一
other h
Tone 2 + 1
其他
outflow l
Tone 2 + 1
Tone 2 + 3
流失
eve l
Tone 2 + 1
前夕
amputation l
Tone 2 + 1
截肢
cabin l
Tone 2 + 1
船舱
piles l
Tone 2 + 1
成堆
echo l
Tone 2 + 1
回声
nothing_more_than l
Tone 2 + 1
无非
outflow l
Tone 2 + 1
流失
violate_regulations l
Tone 2 + 1
违章
chrysanthemun l
Tone 1 + 5
Tone 2 + 1
Tone 2 + 5
Tone 4 + 1
菊花
pitch in z-score
1st syllable
2nd syllable

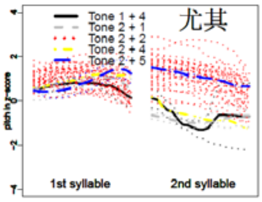
simultaneously h



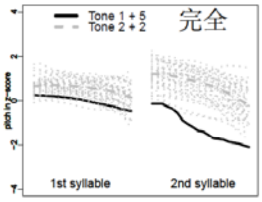
kitchen h



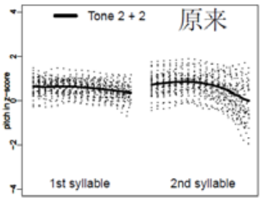
in_particular h



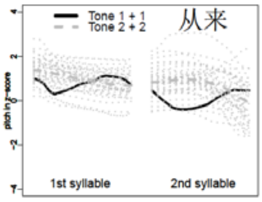
entirely h



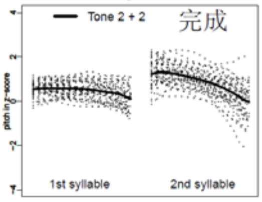
original h



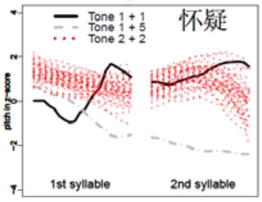
always h



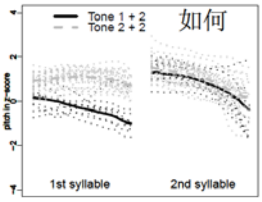
complete h



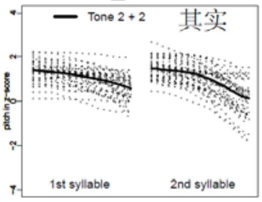
suspect h



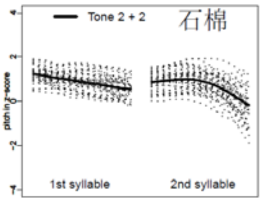
how h



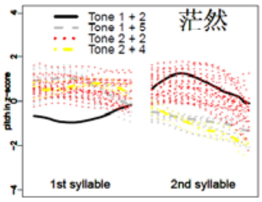
in_fact h



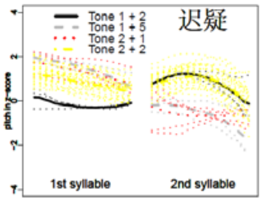
asbestos l



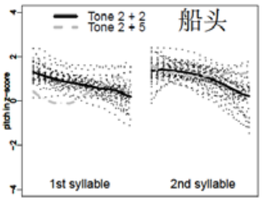
dazed l



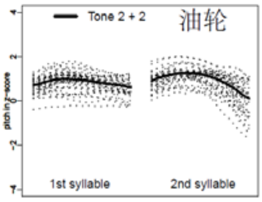
hesitate l

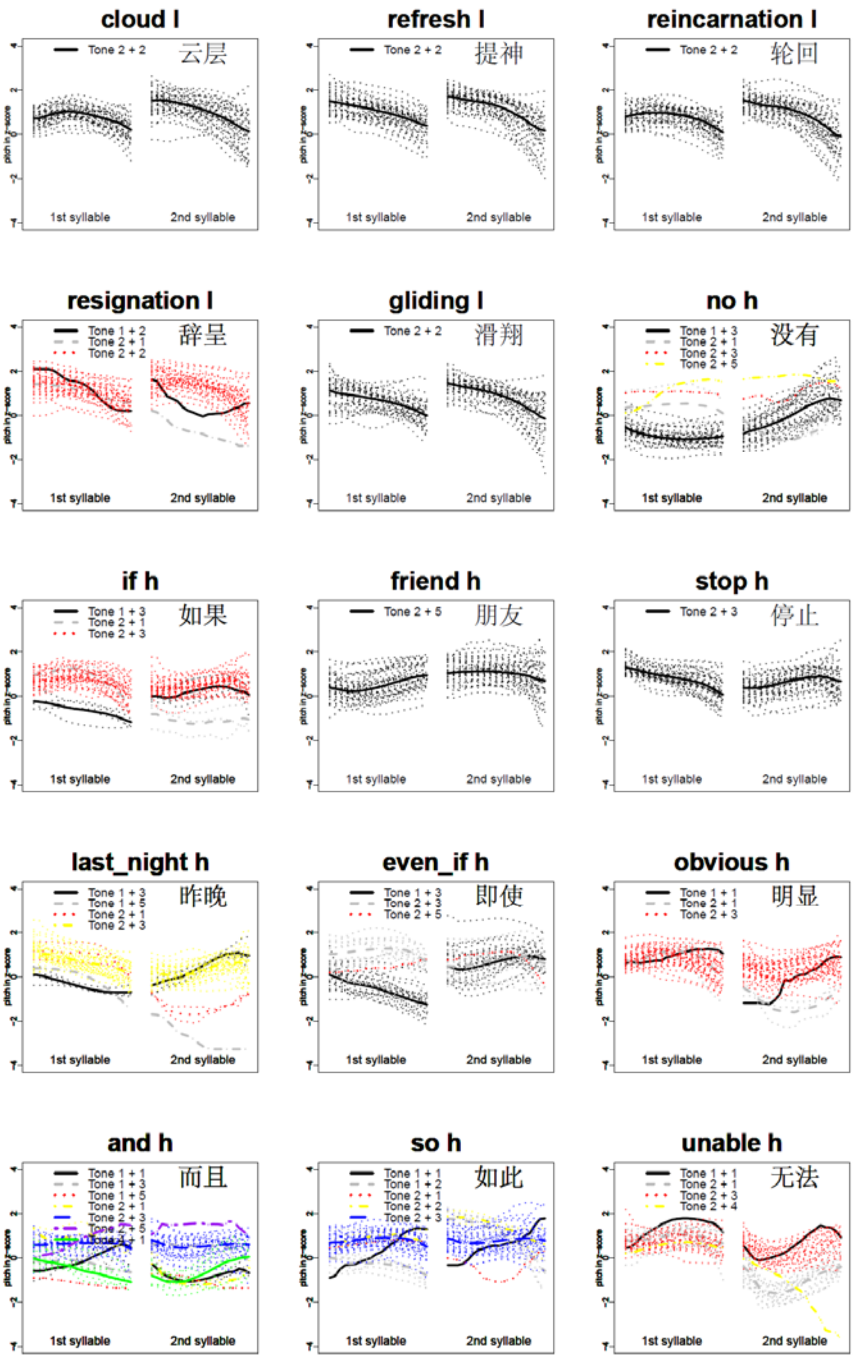


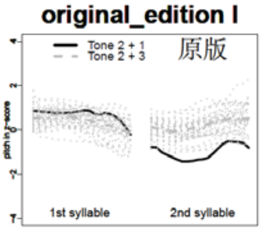
bow_of_the_ship l

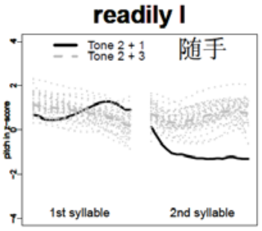


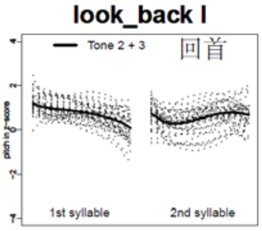
tankers l

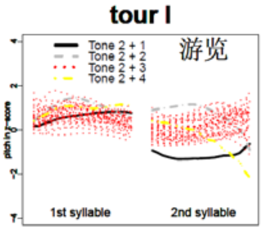


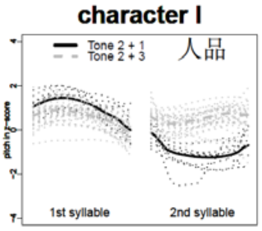

cloud l
Tone 2 + 2
云层
refresh l
Tone 2 + 2
提神
reincarnation l
Tone 2 + 2
轮回
resignation l
Tone 1 + 2
Tone 2 + 1
Tone 2 + 2
辞呈
gliding l
Tone 2 + 2
滑翔
no h
Tone 1 + 3
Tone 2 + 1
Tone 2 + 3
Tone 2 + 5
没有
if h
Tone 1 + 3
Tone 2 + 1
Tone 2 + 3
如果
friend h
Tone 2 + 5
朋友
stop h
Tone 2 + 3
停止
last_night h
Tone 1 + 3
Tone 1 + 5
Tone 2 + 1
Tone 2 + 3
昨晚
even_if h
Tone 1 + 3
Tone 2 + 3
Tone 2 + 5
即使
obvious h
Tone 1 + 1
Tone 2 + 1
Tone 2 + 3
明显
and h
Tone 1 + 1
Tone 1 + 3
Tone 1 + 5
Tone 2 + 1
Tone 2 + 3
Tone 2 + 5
而且
so h
Tone 1 + 1
Tone 1 + 2
Tone 2 + 1
Tone 2 + 2
Tone 2 + 3
如此
unable h
Tone 1 + 1
Tone 2 + 1
Tone 2 + 3
Tone 2 + 4
无法
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

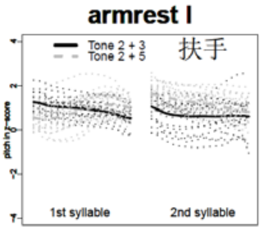

original_edition l
Tone 2 + 1
Tone 2 + 3
原版
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

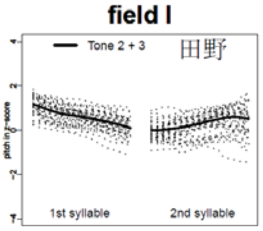

readily l
Tone 2 + 1
Tone 2 + 3
随手
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

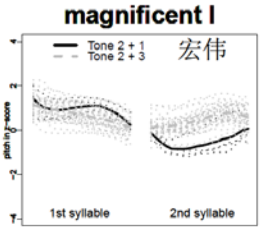

look_back l
Tone 2 + 3
回首
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

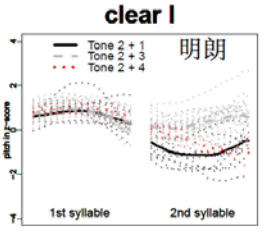

tour l
Tone 2 + 1
Tone 2 + 2
Tone 2 + 3
Tone 2 + 4
游览
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

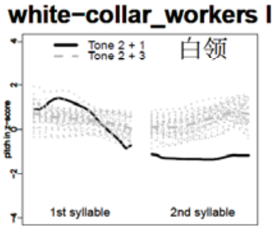

character l
Tone 2 + 1
Tone 2 + 3
人品
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

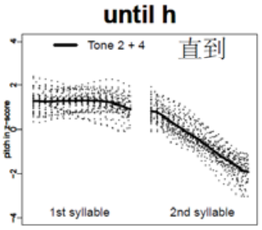

armrest l
Tone 2 + 3
Tone 2 + 5
扶手
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

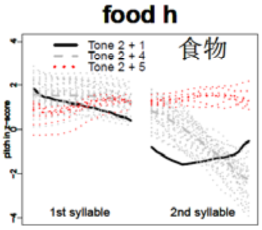

field l
Tone 2 + 3
田野
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

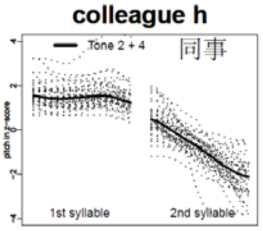

magnificent l
Tone 2 + 1
Tone 2 + 3
宏伟
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

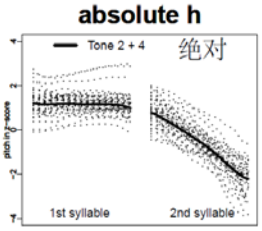

clear l
Tone 2 + 1
Tone 2 + 3
Tone 2 + 4
明朗
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

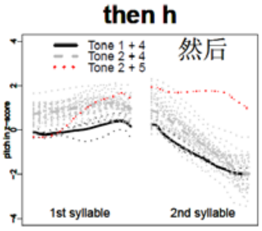

white−collar_workers l
Tone 2 + 1
Tone 2 + 3
白领
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

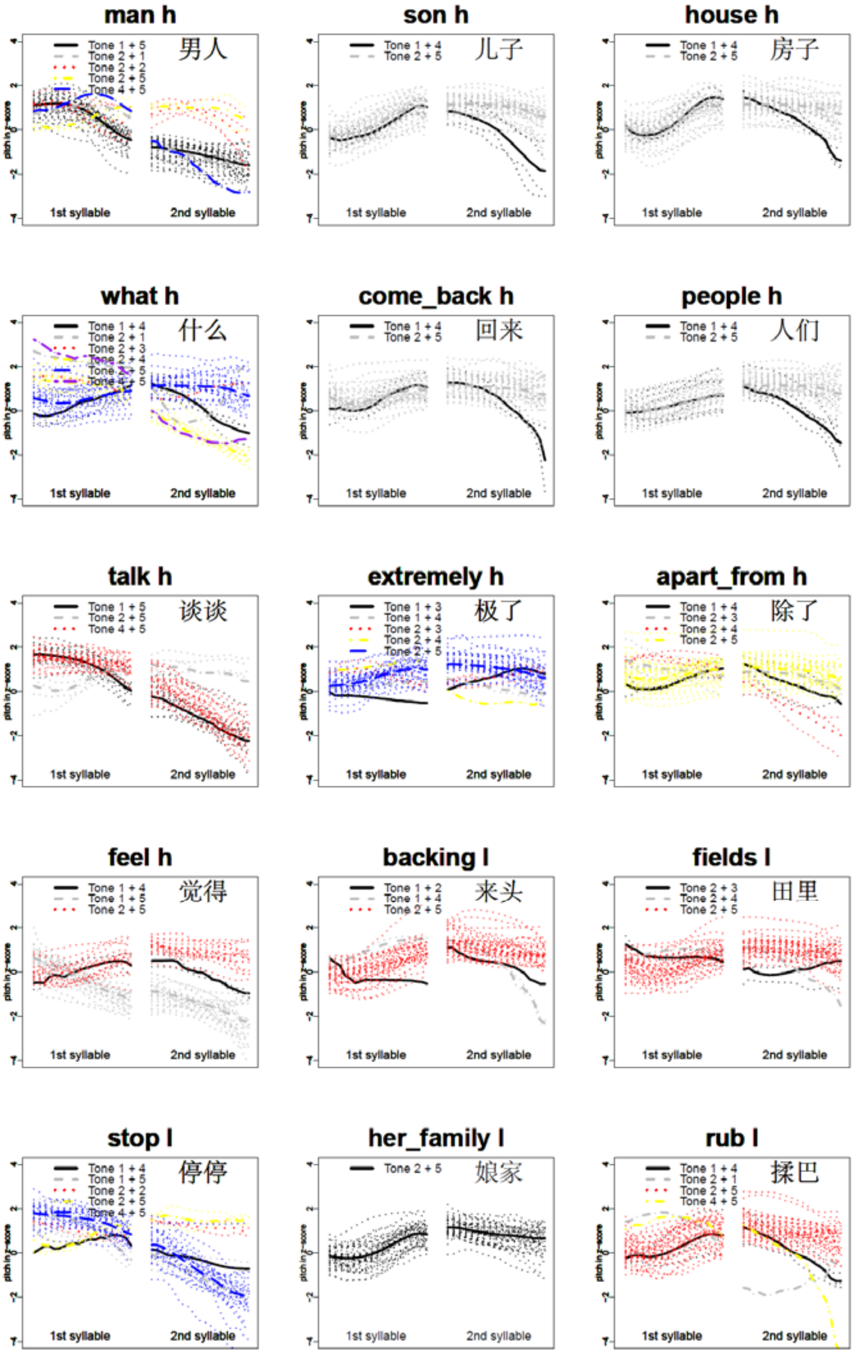

until h
Tone 2 + 4
直到
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

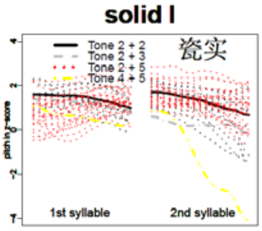

food h
Tone 2 + 1
Tone 2 + 4
Tone 2 + 5
食物
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

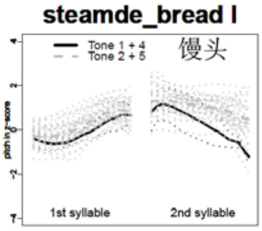

colleague h
Tone 2 + 4
同事
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

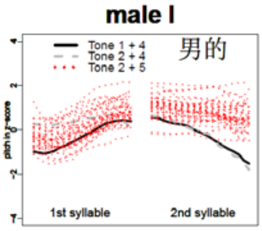

absolute h
Tone 2 + 4
绝对
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

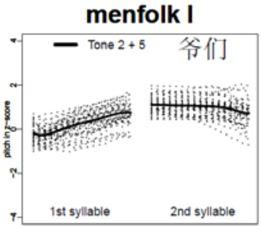

then h
Tone 1 + 4
Tone 2 + 4
Tone 2 + 5
然后
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

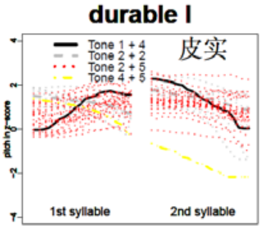

man h
Tone 1 + 5
Tone 2 + 1
Tone 2 + 2
Tone 2 + 5
Tone 4 + 5
男人
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable
son h
Tone 1 + 4
Tone 2 + 5
儿子
house h
Tone 1 + 4
Tone 2 + 5
房子
what h
Tone 1 + 4
Tone 2 + 1
Tone 2 + 3
Tone 2 + 4
Tone 2 + 5
Tone 4 + 5
什么
come_back h
Tone 1 + 4
Tone 2 + 5
回来
people h
Tone 1 + 4
Tone 2 + 5
人们
talk h
Tone 1 + 5
Tone 2 + 5
Tone 4 + 5
谈谈
extremely h
Tone 1 + 3
Tone 1 + 4
Tone 2 + 3
Tone 2 + 4
Tone 2 + 5
极了
apart_from h
Tone 1 + 4
Tone 2 + 3
Tone 2 + 4
Tone 2 + 5
除了
feel h
Tone 1 + 4
Tone 1 + 5
Tone 2 + 5
觉得
backing l
Tone 1 + 2
Tone 1 + 4
Tone 2 + 5
来头
fields l
Tone 2 + 3
Tone 2 + 4
Tone 2 + 5
田里
stop l
Tone 1 + 4
Tone 1 + 5
Tone 2 + 2
Tone 2 + 5
Tone 4 + 5
停停
her_family l
Tone 2 + 5
娘家
rub l
Tone 1 + 4
Tone 2 + 1
Tone 2 + 5
Tone 4 + 5
揉巴

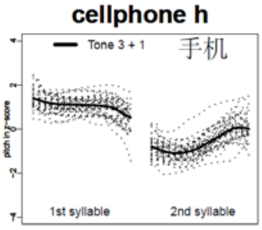

solid l
Tone 2 + 2
Tone 2 + 3
Tone 2 + 5
Tone 4 + 5
瓷实
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

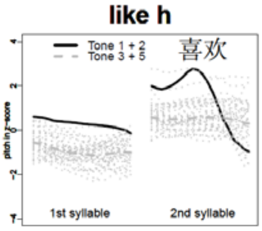

steamde_bread l
Tone 1 + 4
Tone 2 + 5
馒头
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

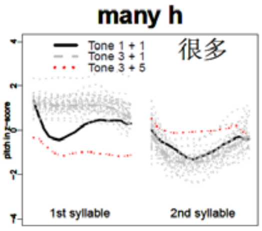

male l
Tone 1 + 4
Tone 2 + 4
Tone 2 + 5
男的
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

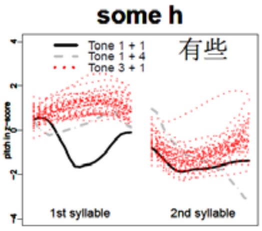

menfolk l
Tone 2 + 5
爷们
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

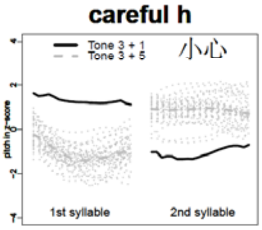

durable l
Tone 1 + 4
Tone 2 + 2
Tone 2 + 5
Tone 4 + 5
皮实
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

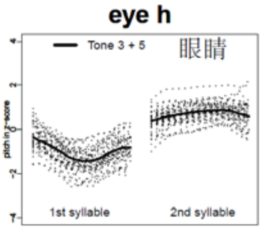

cellphone h
Tone 3 + 1
手机
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

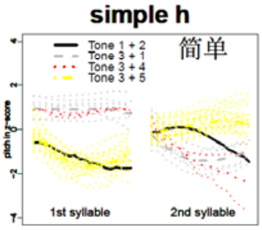

like h
Tone 1 + 2
Tone 3 + 5
喜欢
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

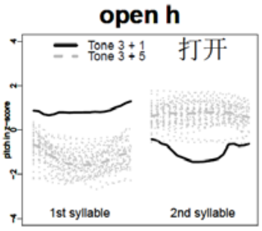

many h
Tone 1 + 1
Tone 3 + 1
Tone 3 + 5
很多
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

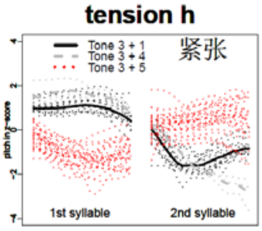

some h
Tone 1 + 1
Tone 1 + 4
Tone 3 + 1
有些
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

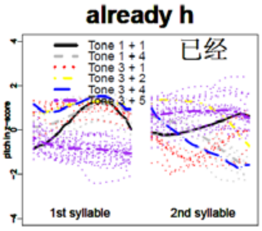

careful h
Tone 3 + 1
Tone 3 + 5
小心
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

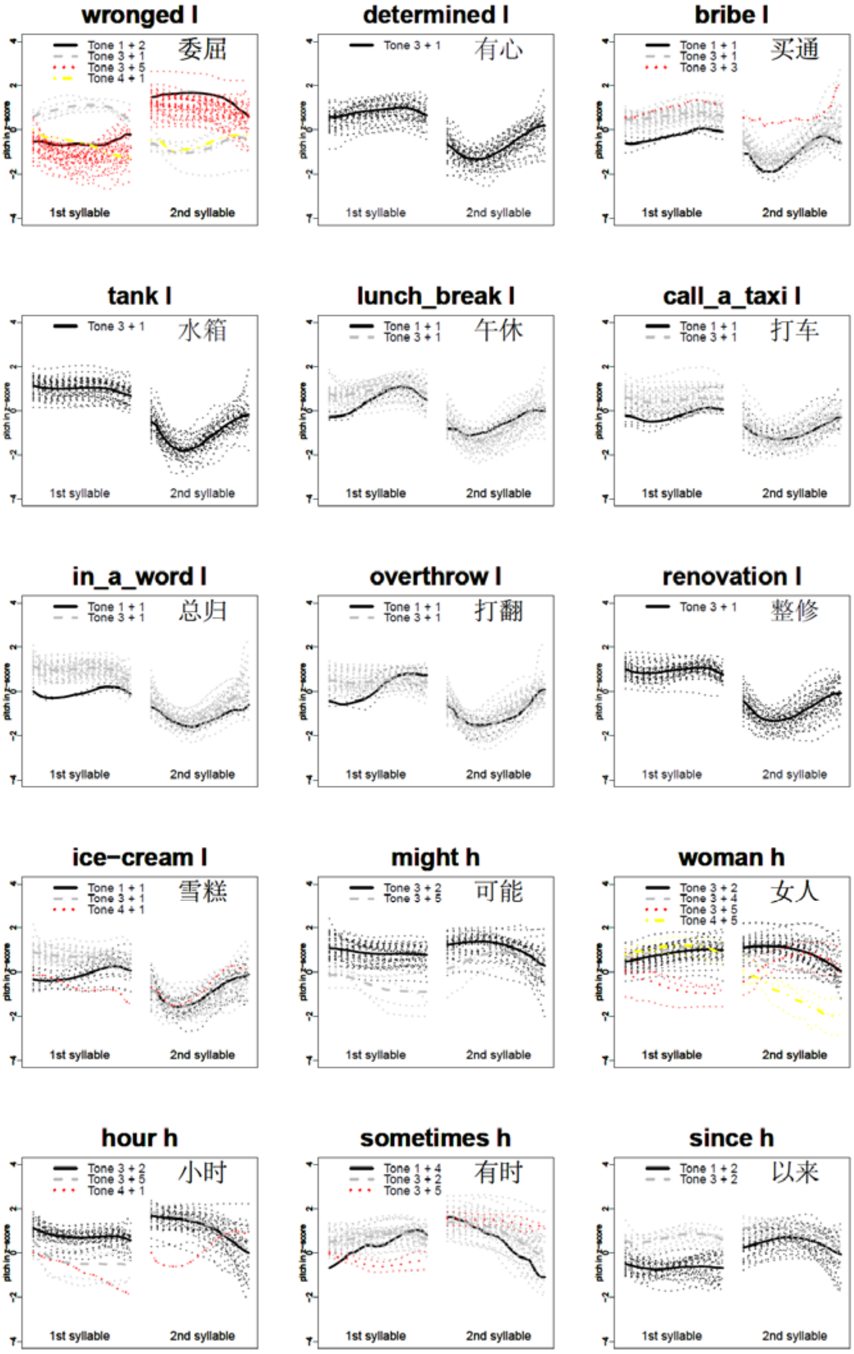

eye h
Tone 3 + 5
眼睛
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

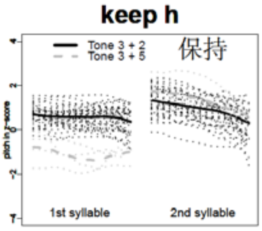

simple h
Tone 1 + 2
Tone 3 + 1
Tone 3 + 4
Tone 3 + 5
简单
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

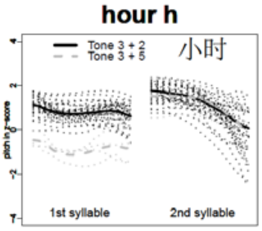

open h
Tone 3 + 1
Tone 3 + 5
打开
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

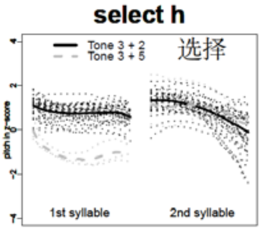

tension h
Tone 3 + 1
Tone 3 + 4
Tone 3 + 5
紧张
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

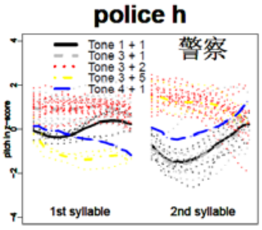

already h
Tone 1 + 1
Tone 1 + 4
Tone 3 + 1
Tone 3 + 2
Tone 3 + 4
Tone 3 + 5
已经
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

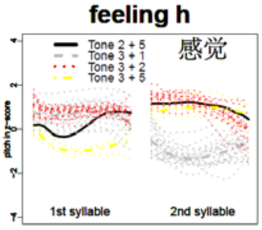

wronged l
Tone 1 + 2
Tone 3 + 1
Tone 3 + 5
Tone 4 + 1
委屈
determined l
Tone 3 + 1
有心
bribe l
Tone 1 + 1
Tone 3 + 1
Tone 3 + 3
买通
tank l
Tone 3 + 1
水箱
lunch_break l
Tone 1 + 1
Tone 3 + 1
午休
call_a_taxi l
Tone 1 + 1
Tone 3 + 1
打车
in_a_word l
Tone 1 + 1
Tone 3 + 1
总归
overthrow l
Tone 1 + 1
Tone 3 + 1
打翻
renovation l
Tone 3 + 1
整修
ice−cream l
Tone 1 + 1
Tone 3 + 1
Tone 4 + 1
雪糕
might h
Tone 3 + 2
Tone 3 + 5
可能
woman h
Tone 3 + 2
Tone 3 + 4
Tone 3 + 5
Tone 4 + 5
女人
hour h
Tone 3 + 2
Tone 3 + 5
Tone 4 + 1
小时
sometimes h
Tone 1 + 4
Tone 3 + 2
Tone 3 + 5
有时
since h
Tone 1 + 2
Tone 3 + 2
以来
pitch in z−score
4
2
0
-2
-4
1st syllable
2nd syllable

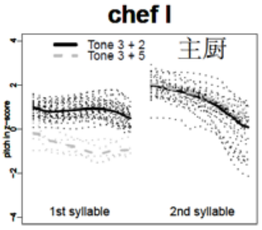

keep h
Tone 3 + 2
Tone 3 + 5
保持
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

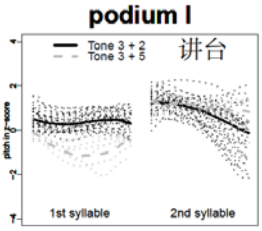

hour h
Tone 3 + 2
Tone 3 + 5
小时
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

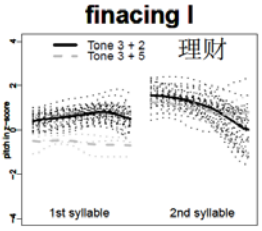

select h
Tone 3 + 2
Tone 3 + 5
选择
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

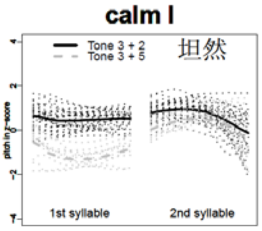

police h
Tone 1 + 1
Tone 3 + 1
Tone 3 + 2
Tone 3 + 5
Tone 4 + 1
警察
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

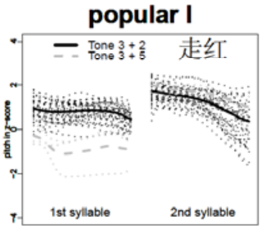

feeling h
Tone 2 + 5
Tone 3 + 1
Tone 3 + 2
Tone 3 + 5
感觉
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

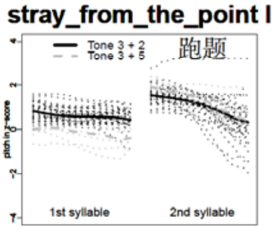

chef l
Tone 3 + 2
Tone 3 + 5
主厨
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

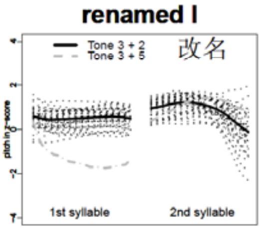

podium l
Tone 3 + 2
Tone 3 + 5
讲台
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

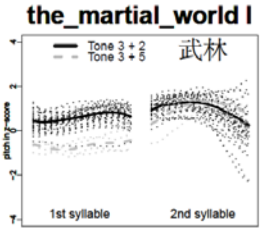

finacing l
Tone 3 + 2
Tone 3 + 5
理财
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

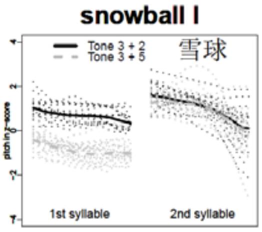

calm l
Tone 3 + 2
Tone 3 + 5
坦然
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

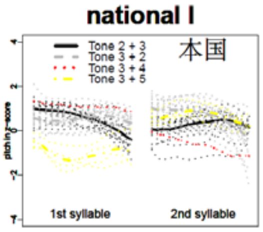

popular l
Tone 3 + 2
Tone 3 + 5
走红
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

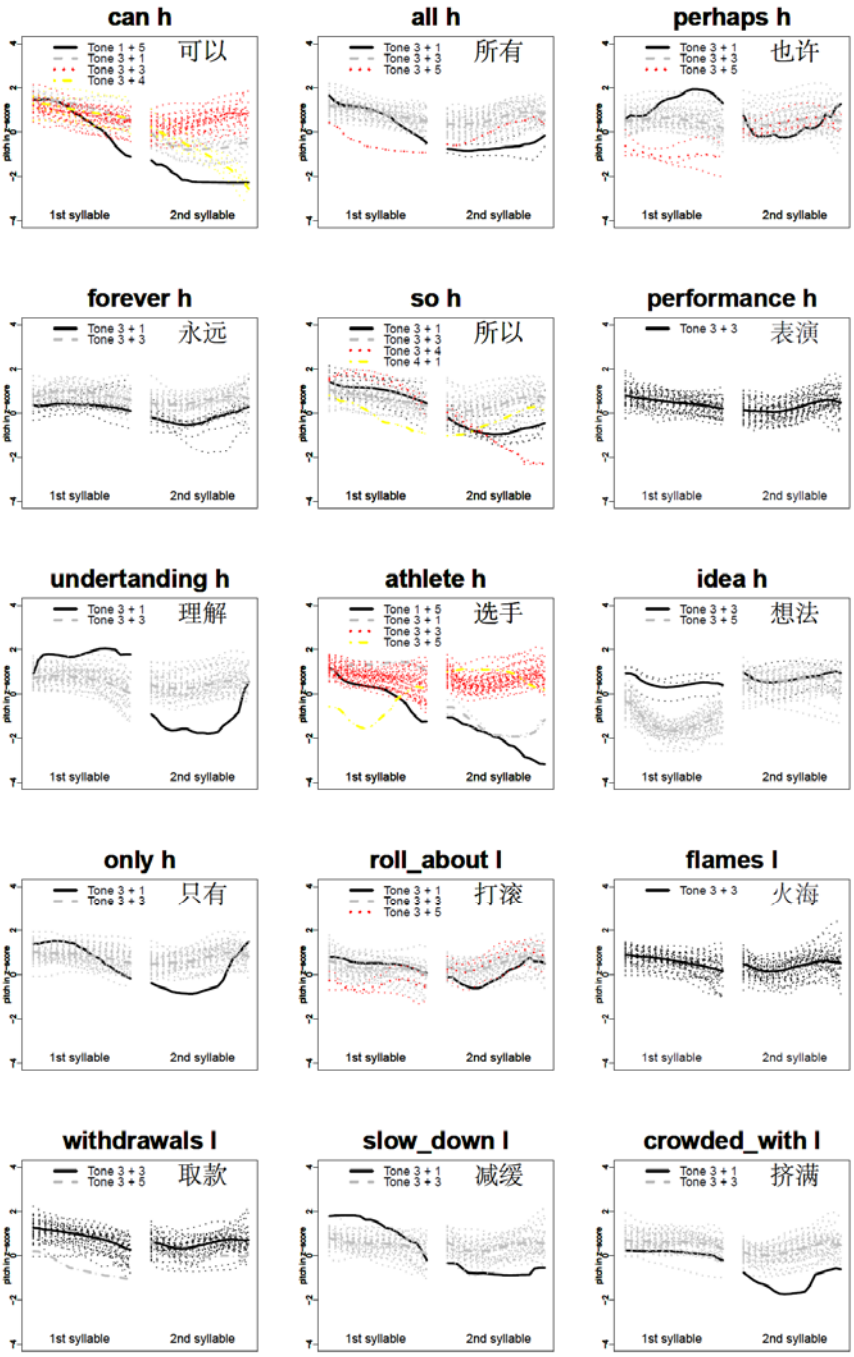

stray_from_the_point l
Tone 3 + 2
Tone 3 + 5
跑题
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

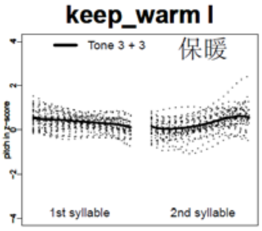

renamed l
Tone 3 + 2
Tone 3 + 5
改名
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

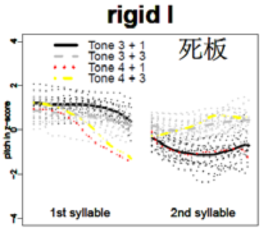

the_martial_world l
Tone 3 + 2
Tone 3 + 5
武林
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

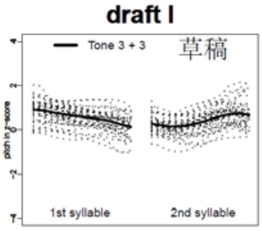

snowball l
Tone 3 + 2
Tone 3 + 5
雪球
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

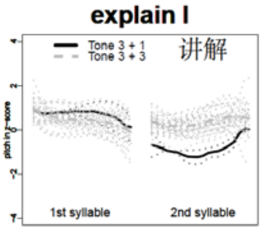

national l
Tone 2 + 3
Tone 3 + 2
Tone 3 + 4
Tone 3 + 5
本国
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

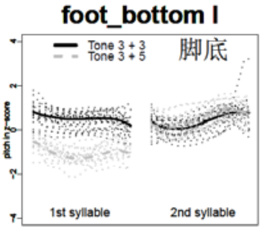

pitch in z-score
4
2
0
-2
-4
can h
Tone 1 + 5
Tone 3 + 1
Tone 3 + 3
Tone 3 + 4
可以
1st syllable
2nd syllable
all h
Tone 3 + 1
Tone 3 + 3
Tone 3 + 5
所有
1st syllable
2nd syllable
perhaps h
Tone 3 + 1
Tone 3 + 3
Tone 3 + 5
也许
1st syllable
2nd syllable
forever h
Tone 3 + 1
Tone 3 + 3
永远
1st syllable
2nd syllable
so h
Tone 3 + 1
Tone 3 + 3
Tone 3 + 4
Tone 4 + 1
所以
1st syllable
2nd syllable
performance h
Tone 3 + 3
表演
1st syllable
2nd syllable
undertanding h
Tone 3 + 1
Tone 3 + 3
理解
1st syllable
2nd syllable
athlete h
Tone 1 + 5
Tone 3 + 1
Tone 3 + 3
Tone 3 + 5
选手
1st syllable
2nd syllable
idea h
Tone 3 + 3
Tone 3 + 5
想法
1st syllable
2nd syllable
only h
Tone 3 + 1
Tone 3 + 3
只有
1st syllable
2nd syllable
roll_about l
Tone 3 + 1
Tone 3 + 3
Tone 3 + 5
打滚
1st syllable
2nd syllable
flames l
Tone 3 + 3
火海
1st syllable
2nd syllable
withdrawals l
Tone 3 + 3
Tone 3 + 5
取款
1st syllable
2nd syllable
slow_down l
Tone 3 + 1
Tone 3 + 3
减缓
1st syllable
2nd syllable
crowded_with l
Tone 3 + 1
Tone 3 + 3
挤满
1st syllable
2nd syllable

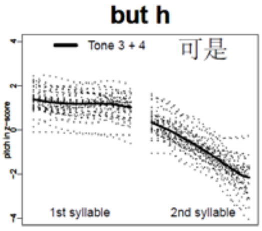

keep_warm l

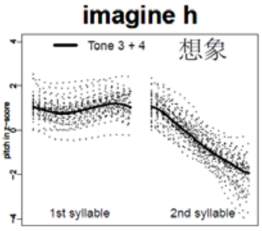

rigid l

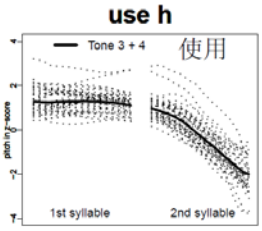

draft l

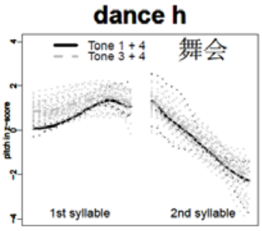

explain l

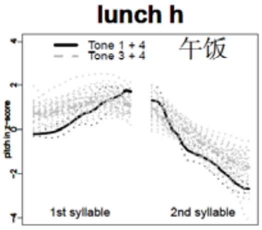

foot_bottom l

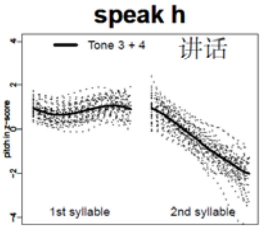

but h

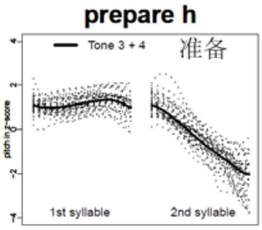

imagine h

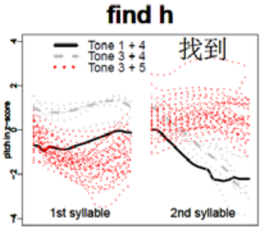

use h

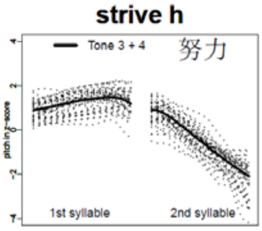

dance h

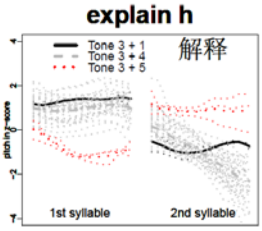

lunch h

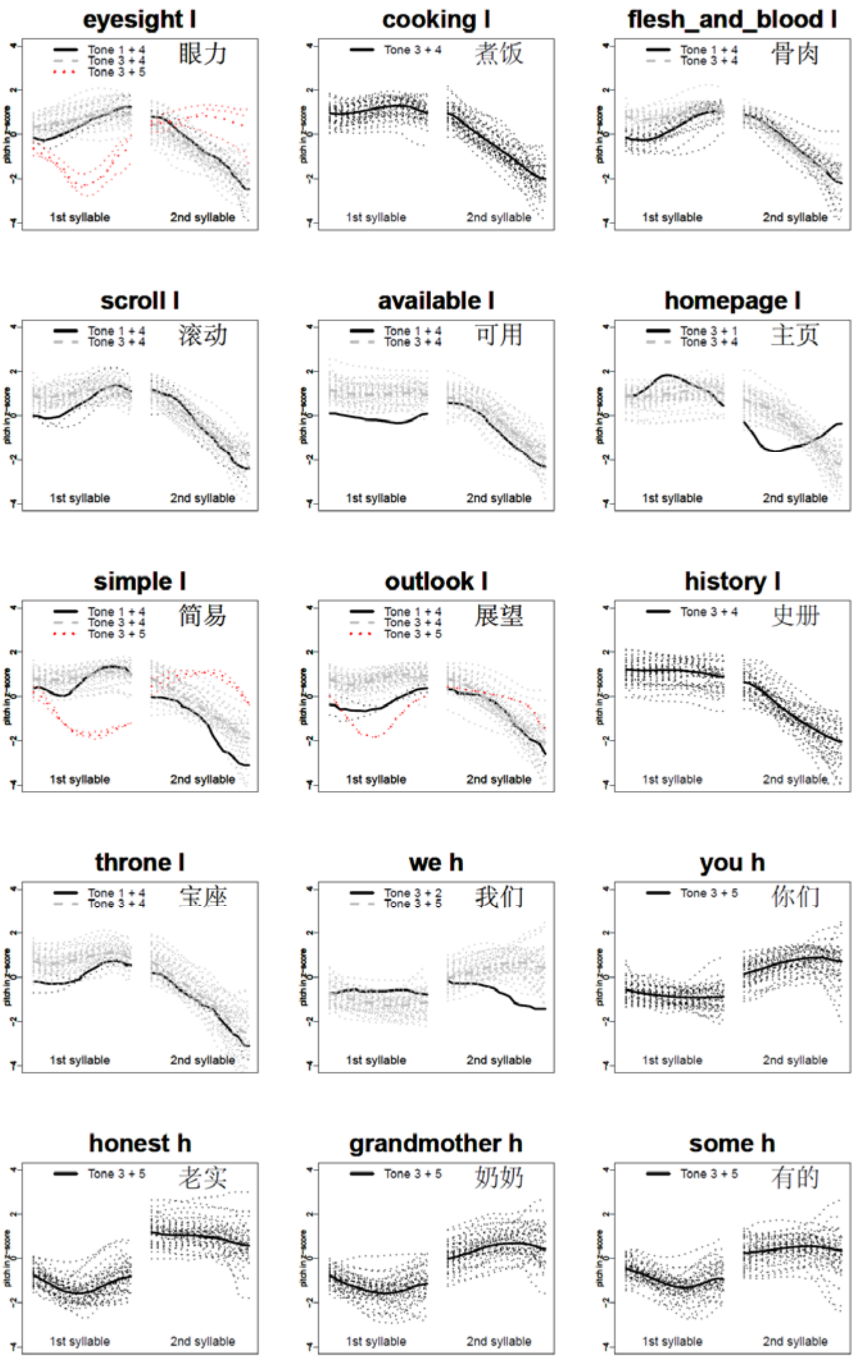

speak h

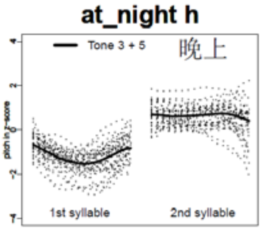

prepare h

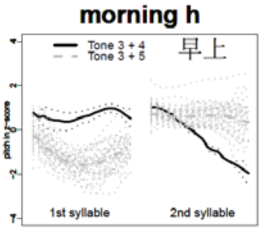

find h

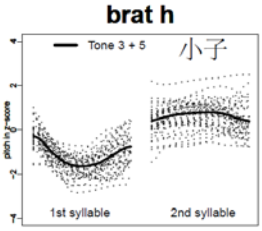

strive h

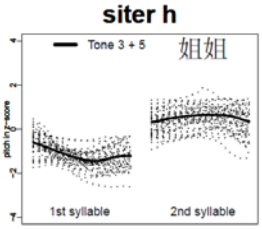

explain h

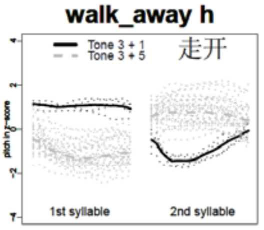

eyesight l
Tone 1 + 4
Tone 3 + 4
Tone 3 + 5
眼力
cooking l
Tone 3 + 4
煮饭
flesh_and_blood l
Tone 1 + 4
Tone 3 + 4
骨肉
scroll l
Tone 1 + 4
Tone 3 + 4
滚动
available l
Tone 1 + 4
Tone 3 + 4
可用
homepage l
Tone 3 + 1
Tone 3 + 4
主页
simple l
Tone 1 + 4
Tone 3 + 4
Tone 3 + 5
简易
outlook l
Tone 1 + 4
Tone 3 + 4
Tone 3 + 5
展望
history l
Tone 3 + 4
史册
throne l
Tone 1 + 4
Tone 3 + 4
宝座
we h
Tone 3 + 2
Tone 3 + 5
我们
you h
Tone 3 + 5
你们
honest h
Tone 3 + 5
老实
grandmother h
Tone 3 + 5
奶奶
some h
Tone 3 + 5
有的
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

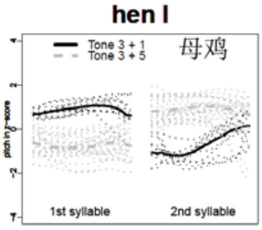

at_night h

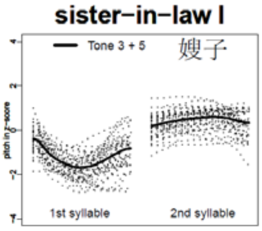

morning h

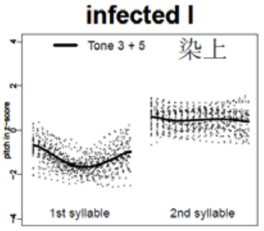

brat h

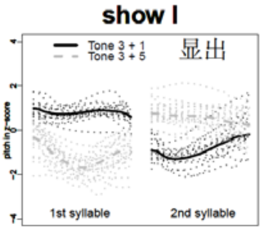

siter h

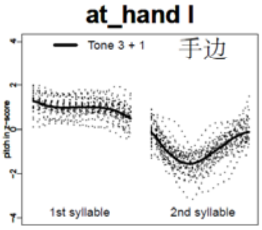

walk_away h

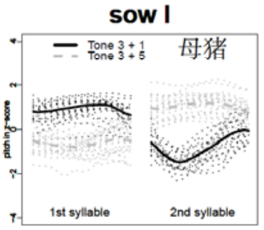

hen l

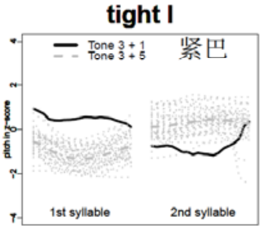

sister−in−law l

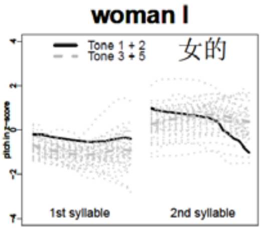

infected l

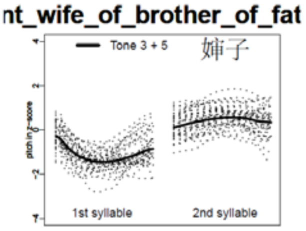

show l

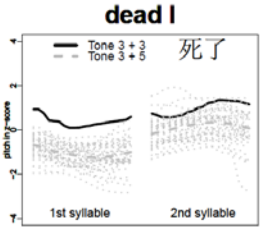

at_hand l

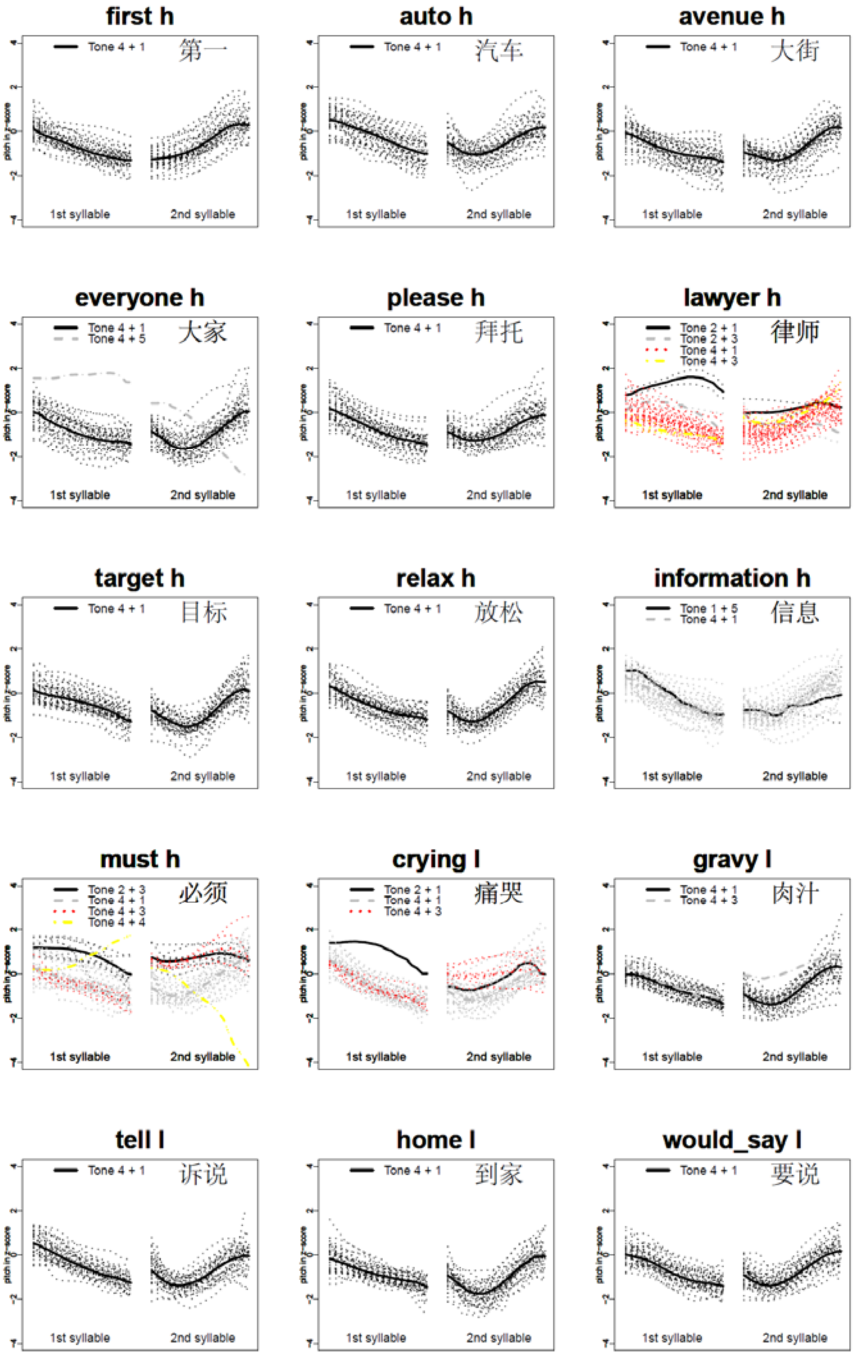

sow l


tight l

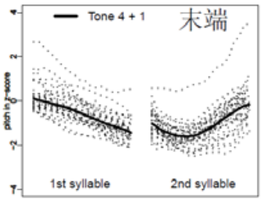

woman l


nt_wife_of_brother_of_fat

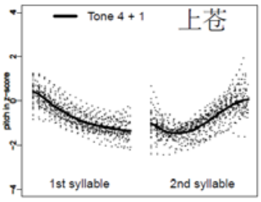

dead l


first h
Tone 4 + 1
第一
auto h
Tone 4 + 1
汽车
avenue h
Tone 4 + 1
大街
everyone h
Tone 4 + 1
Tone 4 + 5
大家
please h
Tone 4 + 1
拜托
lawyer h
Tone 2 + 1
Tone 2 + 3
Tone 4 + 1
Tone 4 + 3
律师
target h
Tone 4 + 1
目标
relax h
Tone 4 + 1
放松
information h
Tone 1 + 5
Tone 4 + 1
信息
must h
Tone 2 + 3
Tone 4 + 1
Tone 4 + 3
Tone 4 + 4
必须
crying l
Tone 2 + 1
Tone 4 + 1
Tone 4 + 3
痛哭
gravy l
Tone 4 + 1
Tone 4 + 3
肉汁
tell l
Tone 4 + 1
诉说
home l
Tone 4 + 1
到家
would_say l
Tone 4 + 1
要说
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

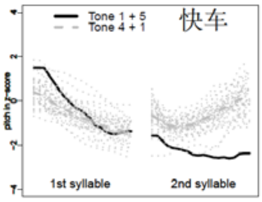
end l



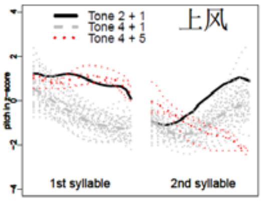
god l



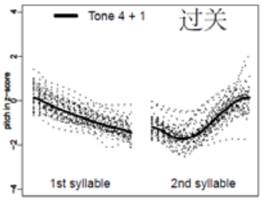
express_train l



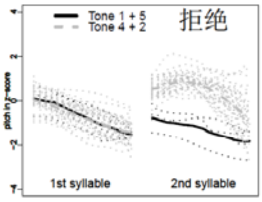
windward l



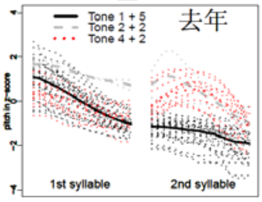
pass l



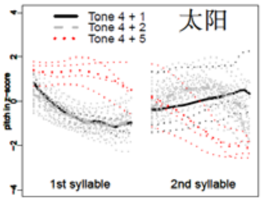
refuse h



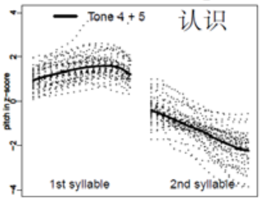
last_year h



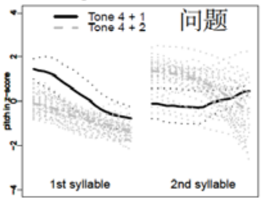
sun h



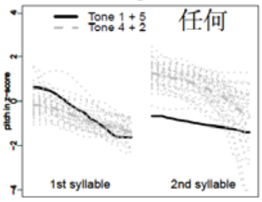
understanding h



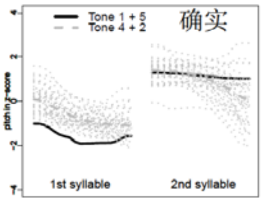
problem h



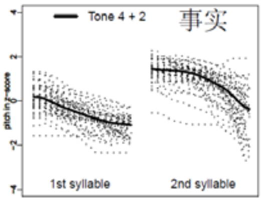
any h



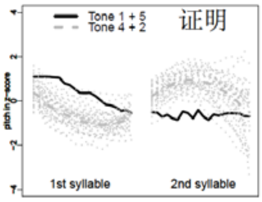
indeed h



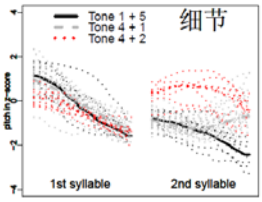
fact h

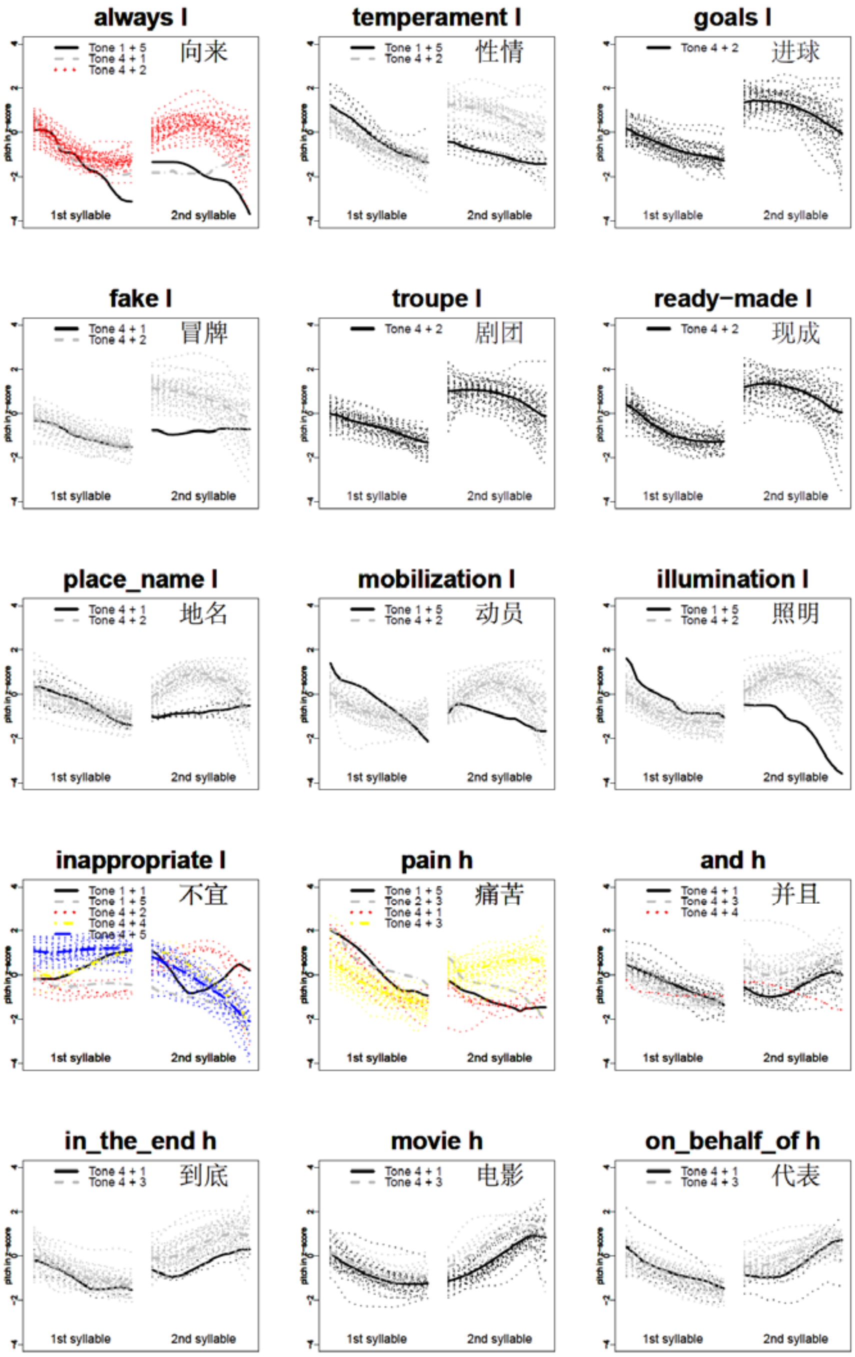


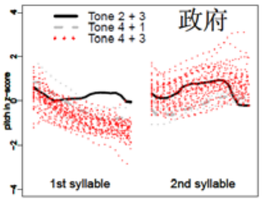
proof h

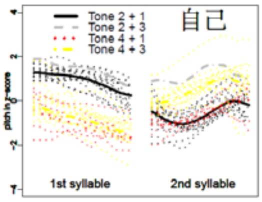


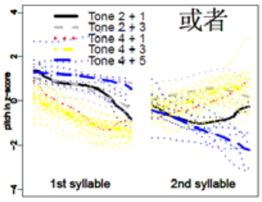
detail h

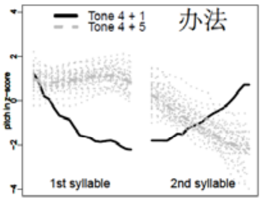


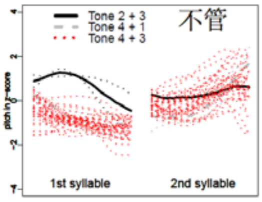

always l
Tone 1 + 5
Tone 4 + 1
Tone 4 + 2
向来
temperament l
Tone 1 + 5
Tone 4 + 2
性情
goals l
Tone 4 + 2
进球
fake l
Tone 4 + 1
Tone 4 + 2
冒牌
troupe l
Tone 4 + 2
剧团
ready−made l
Tone 4 + 2
现成
place_name l
Tone 4 + 1
Tone 4 + 2
地名
mobilization l
Tone 1 + 5
Tone 4 + 2
动员
illumination l
Tone 1 + 5
Tone 4 + 2
照明
inappropriate l
Tone 1 + 1
Tone 1 + 5
Tone 4 + 2
Tone 4 + 4
Tone 4 + 5
不宜
pain h
Tone 1 + 5
Tone 2 + 3
Tone 4 + 1
Tone 4 + 3
痛苦
and h
Tone 4 + 1
Tone 4 + 3
Tone 4 + 4
并且
in_the_end h
Tone 4 + 1
Tone 4 + 3
到底
movie h
Tone 4 + 1
Tone 4 + 3
电影
on_behalf_of h
Tone 4 + 1
Tone 4 + 3
代表
pitch in z−score
4
2
0
-2
-4
1st syllable
2nd syllable

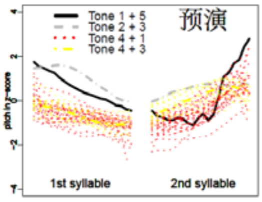

government h

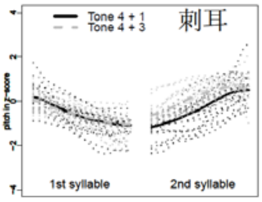

oneself h

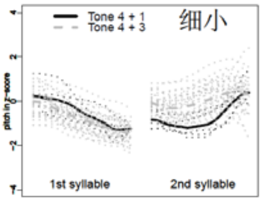

or h

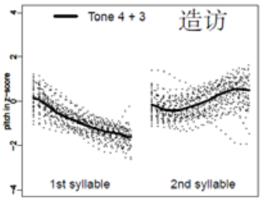

method h

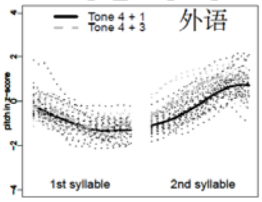

regardless_of h

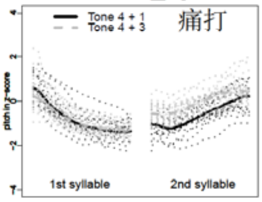

preview l

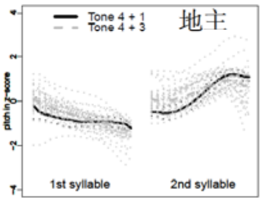

harsh l

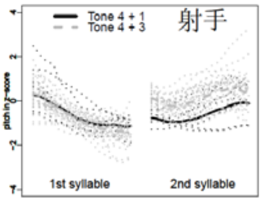

fine l

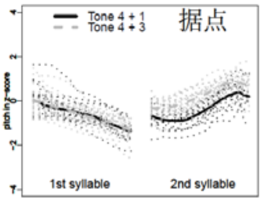

vist l

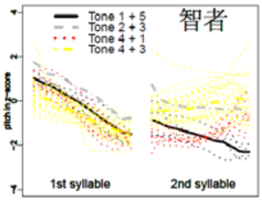

foreign_language l

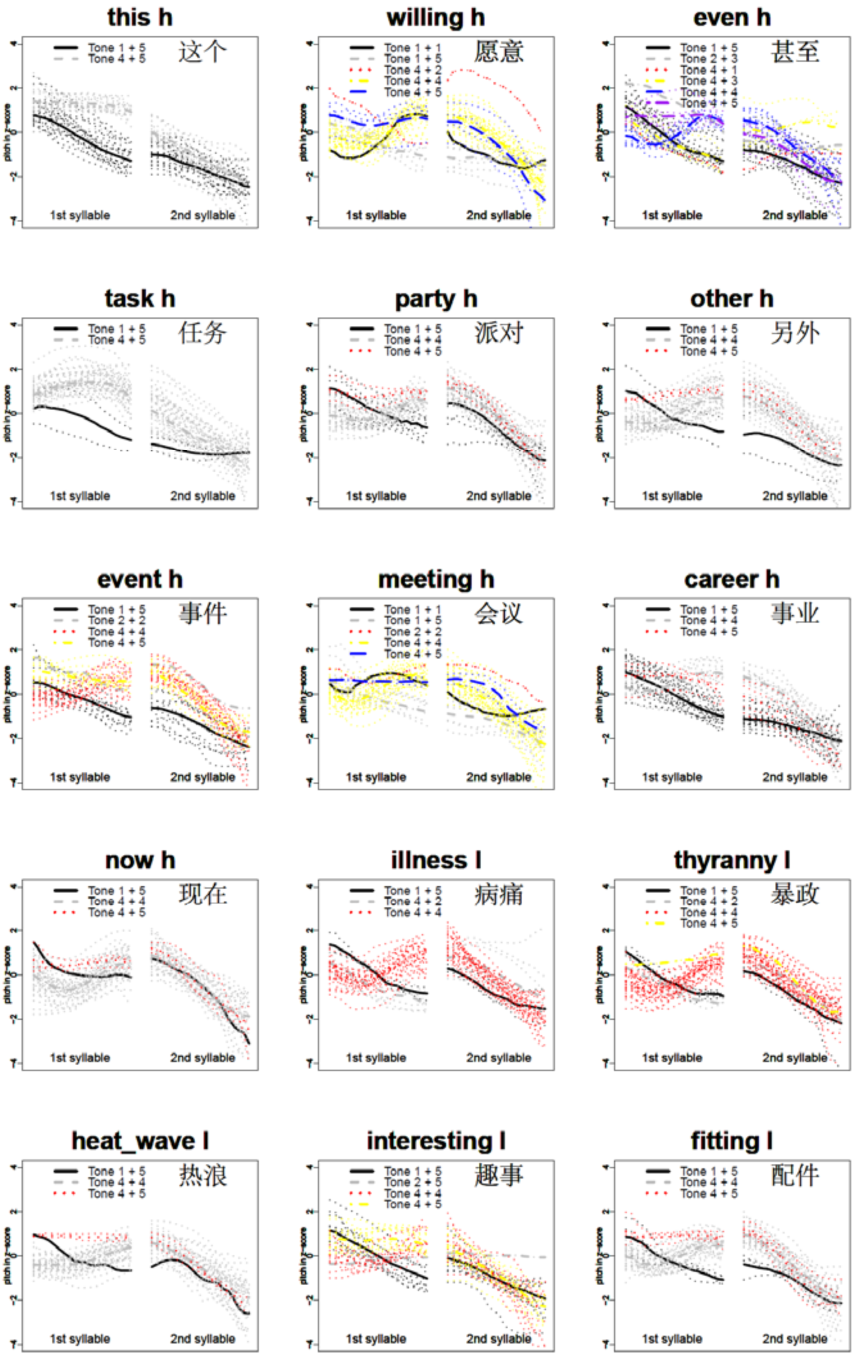

beaten_up l

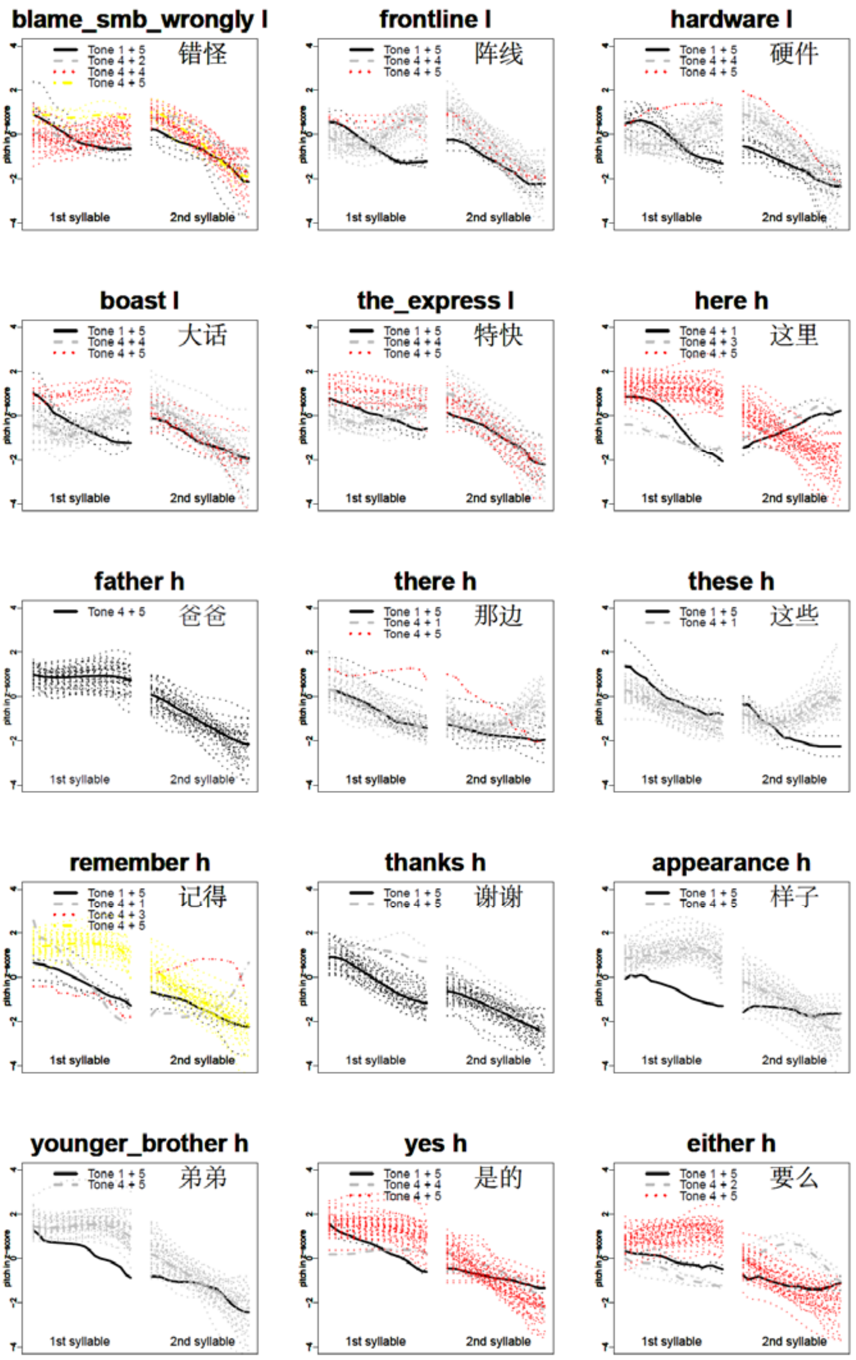

landlords l

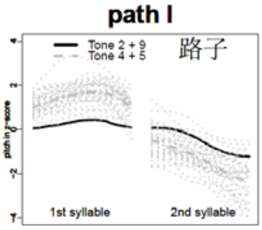

archer l

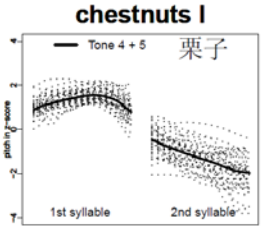

stronghold l

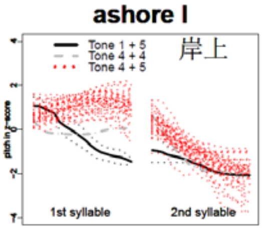

wise_man l

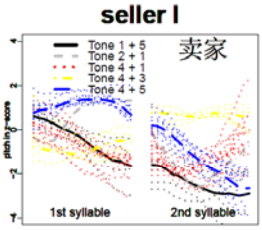

this h
Tone 1 + 5
Tone 4 + 5
这个
willing h
Tone 1 + 1
Tone 1 + 5
Tone 4 + 2
Tone 4 + 4
Tone 4 + 5
愿意
even h
Tone 1 + 5
Tone 2 + 3
Tone 4 + 1
Tone 4 + 3
Tone 4 + 4
Tone 4 + 5
甚至
task h
Tone 1 + 5
Tone 4 + 5
任务
party h
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
派对
other h
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
另外
event h
Tone 1 + 5
Tone 2 + 2
Tone 4 + 4
Tone 4 + 5
事件
meeting h
Tone 1 + 1
Tone 1 + 5
Tone 2 + 2
Tone 4 + 4
Tone 4 + 5
会议
career h
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
事业
now h
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
现在
illness l
Tone 1 + 5
Tone 4 + 2
Tone 4 + 4
病痛
thyranny l
Tone 1 + 5
Tone 4 + 2
Tone 4 + 4
Tone 4 + 5
暴政
heat_wave l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
热浪
interesting l
Tone 1 + 5
Tone 2 + 5
Tone 4 + 4
Tone 4 + 5
趣事
fitting l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
配件
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

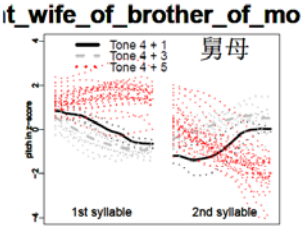

blame_smb_wrongly l
Tone 1 + 5
Tone 4 + 2
Tone 4 + 4
Tone 4 + 5
错怪
frontline l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
阵线
hardware l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
硬件
boast l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
大话
the_express l
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
特快
here h
Tone 4 + 1
Tone 4 + 3
Tone 4 + 5
这里
father h
Tone 4 + 5
爸爸
there h
Tone 1 + 5
Tone 4 + 1
Tone 4 + 5
那边
these h
Tone 1 + 5
Tone 4 + 1
这些
remember h
Tone 1 + 5
Tone 4 + 1
Tone 4 + 3
Tone 4 + 5
记得
thanks h
Tone 1 + 5
Tone 4 + 5
谢谢
appearance h
Tone 1 + 5
Tone 4 + 5
样子
younger_brother h
Tone 1 + 5
Tone 4 + 5
弟弟
yes h
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
是的
either h
Tone 1 + 5
Tone 4 + 2
Tone 4 + 5
要么
pitch in z-score
4
2
0
-2
-4
1st syllable
2nd syllable

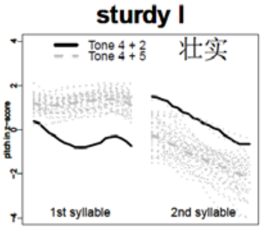

path I
Tone 2 + 9
Tone 4 + 5
路子
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

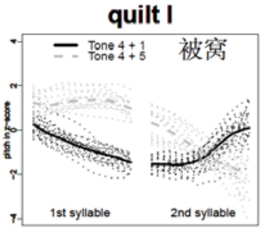

chestnuts I
Tone 4 + 5
栗子
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

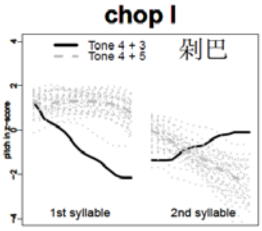

ashore I
Tone 1 + 5
Tone 4 + 4
Tone 4 + 5
岸上
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

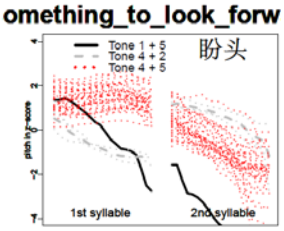

seller I
Tone 1 + 5
Tone 2 + 1
Tone 4 + 1
Tone 4 + 3
Tone 4 + 5
卖家
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

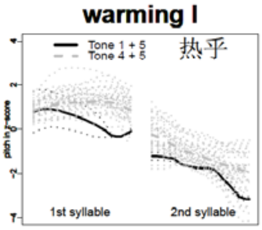

t_wife_of_brother_of_mo
Tone 4 + 1
Tone 4 + 3
Tone 4 + 5
舅母
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable


sturdy I
Tone 4 + 2
Tone 4 + 5
壮实
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable


quilt I
Tone 4 + 1
Tone 4 + 5
被窝
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable


chop I
Tone 4 + 3
Tone 4 + 5
剁巴
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable


omething_to_look_forw
Tone 1 + 5
Tone 4 + 2
Tone 4 + 5
盼头
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable


warming I
Tone 1 + 5
Tone 4 + 5
热乎
4
2
0
-2
-4
pitch in z-score
1st syllable
2nd syllable

Nederlandse Samenvatting

Dit proefschrift onderzoekt het lexicale verwerkingssysteem bij een speciaal soort tweetaligheid, dat gekenmerkt wordt door (1) twee toonsystemen, (2) twee nauw verwante dialecten en (3) het bezit van een gemeenschappelijk logografisch schrift.

Hoewel de rol van de lexicale tonen bij lexicale toegang steeds meer aandacht krijgt, is nog altijd onvoldoende bekend hoe twee toonsystemen op elkaar inwerken in het tweetalige mentale lexicon. Daarnaast weten historisch taalkundigen al lang dat er systematische correspondenties bestaan in de woordenschat van nauw verwante talen en dialecten. Zulke systematische correspondenties zijn echter zelden in verband gebracht met de rijkdom aan lexicale varianten in de woordproductie door tweetaligen. Die is bij tweetalige individuen ook niet eerder onderzocht vanuit sociolinguïstisch en cognitief perspectief. Bovendien veronderstelden de meeste eerdere onderzoeken naar tweetalige visuele woordherkenning dat de tweetaligen gebruik maakten van ofwel hetzelfde type alfabetische schrift, ofwel van een per taal verschillend schriftsysteem. De mogelijkheid dat tweetaligen in twee nauw verwante dialecten gebruik maken van hetzelfde logografische werd daarbij niet onderzocht. In dat geval kan, ondanks het fonologische verschil tussen de etymologisch verwante vertaalequivalenten, exact dezelfde visuele woordvorm worden gebruikt. Hoe dergelijke gemeenschappelijke visuele woordvormen fonologische informatie in de mentale lexicon van de tweetaligen activeren vergt nader onderzoek.

Standaard Chinees (SC) en Jinan Mandarijn (JM) zijn beide Mandarijndialecten. SC is de officiële standaardtaal van China, waarvan de fonologie gebaseerd is op het dialect van Beijing. JM is het Mandarijndialect gesproken in Jinan, 497 km ten oosten van Beijing. Met tweetaligheid in SC en JM als testcase, wordt in deze studie een reeks van theoretische en empirische problemen bezien, met speciale aandacht voor de rol van de tooninformatie bij de lexicale toegang. Om de bijbehorende onderzoeksvragen te beantwoorden, werden met behulp van gedragsexperimenten gegevens verzameld op gebied van spraakproductie en van woordbegrip. Aangezien dit proefschrift op een reeks artikelen is gebaseerd, zijn de meer gedetailleerde onderzoekachtergronden, bespreking van standpunten en resultaten alsook de conclusies binnen elk hoofdstuk apart te vinden.

De acht hoofdstukken worden hieronder samengevat.

Hoofdstuk 1 is de algemene inleiding. Na een korte introductie van de fundamentele onderzoeksvragen, wordt een overzicht gegeven voor elk van de samenhangende onderzoeksgebieden. Het overzicht is georganiseerd rond de drie bijzondere aspecten van de specifieke tweetaligheid die het onderwerp vormt van dit proefschrift. Ten eerste, heeft deze tweetaligheid betrekking op alleen de toonsystemen. Een overzicht wordt gegeven op het bijzondere van toontalen in spraakperceptie en lexicale toegang. Vervolgens besteed ik aandacht aan het perceptuele leren van niet-inheemse woordtonen, aan interlinguale toonwaarneming, en aan andere tweetaligheidsstudies met betrekking tot het Chinees. Ten tweede, heeft deze tweetaligheid betrekking op nauw verwante dialecten. Anders dan bij nietverwante talen, vertonen nauw verwante dialecten zowel overeenkomst in hun fonologische inventarissen alsook systematische correpondentie in hun woordenschat. Zo zijn de tonen in eenlettergrepige JM woorden in hoge mate voorspelbaar uit de woordtoon die zij hebben in hun vertaalequivalent in SC. Omdat systematische correpondentie weinig aandacht heeft gekregen in het tweetaligheidsonderzoek, geef ik in deze paragraaf een korte inleidende behandeling van het mechanisme en het corresponderende toonverschijnselen in de onderzochte tweetaligheid. Ten derde, speelt in dit geval van tweetaligheid het bezit van een gemeenschappelijk logografisch schriftsysteem een belangrijke rol. Ik geef daarom een overzicht van wat bekend is over hoe in het leesproces bij tweetaligen woorden in de ene taal hun pendanten in de andere taal kunnen activeren. Na deze algemene overzichten, volgen aparte theoretische inleidingen op elke individueel hoofdstuk.

Hoofdstuk 2 onderzoekt een geval van een twee-op-een afbeelding tussen de tonen in twee toonsystemen. Dit hoofdstuk richt zich op de vraag in hoeverre de sterkte van de vormovereenkomst tussen de woordtonen in de verschillende toonsystemen zijn effect behoudt op de lexicale toegang en tekstbegrip in de andere taal. Dat er een asymmetrische twee-op-een afbeelding tussen de stijgende toon in JM en de hoge en lage stijgende tonen in SC bestaat, wordt geverifieerd aan de hand van de verdeling van hun akoestische eigenschappen en van hun waarneming in de andere taal. Een semantisch primingexperiment toont namelijk aan dat interlinguale gelijkenis tussen woordtonen alleen van invloed is op lexicale toegang maar geen effect heeft op de semantische activering na lexicale retrieval. Dit wijst erop dat deze twee fasen in het spraakverwerkingsproces tot op zekere hoogte gescheiden zijn.

Hoofdstuk 3 doet statistisch onderzoek naar de effecten van systematische tooncorrespondenties in de woordenschat van SC en JM in bij de spraakproductie. Dit hoofdstuk richt zich op de vraag hoe de sterkte van systemantische correspondentie wordt beïnvloed door de sociolinguïstische en cognitieve achtergrond van de proefpersoon. De toonhoogte-afstand tussen de woorden in JM kan worden voorspeld uit hun tooncategorieën in SC met behulp van statistische modellen, als interlinguale toonidentiteit en de individuele achtergrond van de proefpersoon in aanmerking worden genomen. Deze globale interactie tussen de sociolinguïstische en cognitieve achtergrond van de individuele tweetalige enerzijds en de sterkte van de systematische correspondentie anderzijds is in eerder onderzoek nooit aangetroffen. Anders dan de andere hoofdstukken, kent dit hoofdstuk een meer interdisciplinaire benadering. Ik bespreek de bevindingen in dit hoofdstuk in het licht van de theorieën van taalverandering en cognitieve veroudering bij tweetaligen. Ook bespreek ik de implicaties van de resultaten voor de spraaktechnologie.

Hoofdstuk 4 onderzoekt verder het effect van toonovereenkomst tussen etymologisch verwante vertaalequivalenten op de auditieve lexicale toegang. Tonale overeenkomst heeft discontinue (faciliterende dan wel bemoeilijkende) effecten op de lexicale toegang van SC-JM etymologisch verwante vertaalequivalenten en interageert ook op een complexe manier met de taalmodus (d.w.z. JM of SC) en proefblok. Uit het feit echter dat het effect van taaldominatie (d.w.z. met SC als dominant dialect wordt een SC woord sneller verwerkt dan zijn JM vertaalequivalent) overeind blijft volgt dat de lexicale representaties van de vertaalequivalenten in het tweetalig lexicon zich niet alleen onderscheiden in de woordtonen maar ook door de taaldominantie. Bovendien hebben de SC-JM tweetaligen een bijzonder lexicaal voordeel ten opzichte van de eentalige controleproefpersonen. Dit geeft aan dat tweetaligen met nauwverwante dialecten, met een mentaal lexicon vol etymologisch verwante vertaalequivalenten, voordeel ondervinden bij de auditieve lexicale toegang.

In JM vinden we een aanzienlijk aantal lexicale toonvarianten in de spraakproductie; die worden onderzocht in hoofdstuk 5 en 6. Deze twee hoofdstukken probeer ik antwoord te geven op de vraag hoe lexicale toonvarianten liggen opgeslagen en toegangelijk zijn in het tweetalig lexicon.

Hoofdstuk 5 onderzoekt de rol van verschillende soorten toonvariatie bij de auditieve lexicale toegang in JM, met speciale aandacht voor de verwerking van patroonvariatie tussen lexicale toonvarianten. Effecten van exacte (token) herhaling, variatie binnen de categorie, toonpatroonvariatie, en lexicaal contrastieve variabiliteit worden vergeleken in een auditief vorm-primingexperiment. De resultaten geven steun aan de opvatting dat meerdere toonpatronen deel kunnen uitmaken van de lexicale representatie van een woord bij de lexicale toegang, maar tegelijkertijd ook convergeren naar één specifieke lexicale vorm.

Hoofdstuk 6 vergelijkt de lexicale toegang bij de spraakproductie van JM lexicale varianten die identiek of juist niet-identiek zijn met hun SC vertaalequivalenten. Uit de resultaten blijkt een frequentie-effect dat aangeeft dat SC-JM tweetaligen (hoogfrequente) JM woordtoonvarianten in hun mentale lexicon opgeslagen hebben liggen, ook als ze die niet zelf in het corpus hebben produceerd. Het verschil in eenmalige individuele keuzes weerspiegelt dan voornamelijk de relatieve frequentie van de opgeslagen varianten en betreft niet een wezenlijk verschil in de lexicale representaties bij individuele proefpersonen.

Hoofdstuk 7 onderzoekt, aan de hand van een Stroop-taak, de automatische tweetalige fonologische activatie vanuit de gemeenschappelijke geschreven woordvormen, met speciale aandacht voor de vraag hoe de SC-JM (tonale) tweetaligen verschillen van SC (tonale) eentaligen bij het automatisch ophalen van tooninformatie uit Chinese karakters (schrifttekens). Ook al is de (via het schrift uitgelokte) woordproductie bij de tweetaligen over de hele linie zwakker, toch hebben zij voordeel van de congruente condities en ondervinden zij minder last van de incongruente condities dan de eentalige controlepersonen. De SC-JM tweetaligen verschillen ook van de SC (tonale) eentaligen door een gebrek aan toongevoeligheid, dat aan de tweetalige aandachtcontrole kan worden gerelateerd en dat voordelig is bij het uitvoeren van een Stroop-taak.

Hoofdstuk 8, ten slotte, geeft een overzicht van de belangrijkste bevindingen van dit proefschrift en bespreekt deze in samenhang. De bevindingen worden geherstructureerd rond de belangrijkste theoretische vraagstukken. De rol van systematische correspondentie in taalverandering en de invloed van cognitieve veroudering op de structuur van het lexicon van tonale tweetaligen worden besproken aan de hand van de resultaten uit hoofdstuk 3. De resultaten van hoofdstukken 2 en 4 worden tezamen geëvalueerd bij de bespreking van sw rol van interlinguale tonale gelijkenis bij tweetalige lexicale toegang. Ook de hoofdstukken 5 en 6 onderzoeken tonale lexicale varianten met toonpatroonvariatie. De resultaten van deze twee hoofdstukken worden samen besproken, met speciale aandacht voor de lexicale opslag van tonale lexicale varianten. Hoofdstuk 7 geeft nader inzicht in tweetalige lexicale toegang, maar betrekt daar nu ook de automatische fonologische verwerking bij de visuele woorderkenning bij. Als laatste onderdeel worden de tweetalige aandachtsverdeling en executieve controle, evenals de invloed daarvan op de tweetalige lexicale toegang besproken aan de hand van de resultaten van de hoofdstukken 4 en 7.

English Summary

This thesis investigates the lexical processing system in a special type of bilingualism, which involves (1) two tonal systems, (2) two closely related dialects, and (3) a common logographic writing system.

Although the role of lexical tones in lexical access is receiving more and more attention, how two tonal systems interact in the bilingual mental lexicon is largely unknown. Additionally, the fact that systematic correspondence exists between two vocabularies of closely related dialects has long been known by historical linguists. However, the systematic correspondence has rarely been associated with the rich lexical variants in bilingual lexical production. Neither has it been associated with the bilingual individuals' sociolinguistic and cognitive backgrounds. Moreover, most previous studies of bilinguals' visual word recognition assumed that the bilinguals either use similar alphabetic writing systems or different types of writing systems for their two languages. However, the possibility that the same logographic writing system can be used in the bilingualism of two closely related dialects was largely neglected. In this case, in spite of the phonological difference between the etymologically related translation equivalents, exactly the same visual word form can be used. How such common visual word forms activate phonological information in the bilinguals' mental lexicon also needs more investigation.

Standard Chinese (SC) and Jinan Mandarin (JM) are both Mandarin dialects. SC is the official standard language used by China, with the phonology based on the Beijing dialect. JM is the Mandarin dialect spoken in Jinan, 497 km to the east of Beijing. Using the bilingualism of SC and JM as the test case, a series of theoretical and empirical issues are examined in this study, with a focus on the role of tonal information in lexical access. To answer the corresponding research questions, both speech production and lexical comprehension data were collected in the field, using behavioral experiments. Since this thesis is based on a collection of articles the detailed research backgrounds, discussions, and conclusions can be found within each chapter. In the following parts of this summary, a brief description is given for each chapter.

Chapter 1 is the general introduction. After a brief introduction of the basic research questions, each of the related research fields is reviewed. The overview is organized around the three unique aspects of this bilingualism. First, this bilingualism involves two tonal systems. A summary is given on the uniqueness of tonal languages in speech perception and lexical access. Then the perceptual learning of non-native lexical tones, interlingual tonal perception, and other bilingual studies involving Chinese are overviewed. Second, this bilingualism involves two closely related dialects. Different from remote languages, closely related dialects show both similarity between the phonological inventories, and systematic correspondence between the vocabularies. For instance, the tonal categories of monosyllabic JM words are, to a large extent, predictable from the tonal categories of their SC translation equivalents. Since systematic correspondence has received little attention in the research of bilingualism, this session only provides a brief introduction to the mechanism and the corresponding phenomenon in the bilingualism under study. Third, this bilingualism involves a common

logographic writing system. The literature on interlingual activation in bilingual reading is reviewed and the situation in the bilingualism under study is explained with examples. After these general overviews, a separate theoretical introduction is given to each individual main chapter.

Chapter 2 investigates a case of two-to-one tonal mapping between the two tonal systems. This chapter focuses on the extent to which the interlingual category-goodness between different tonal systems keeps its impact on lexical access and speech comprehension. The two-to-one asymmetrical mapping of the JM rising tone and the SC high- and low-rising tones is verified in the distributions of acoustic properties and in interlingual perception. A semantic priming experiment shows that interlingual category-goodness only affects lexical access but does not affect the semantic activation after lexical retrieval, suggesting some discreteness between these two steps.

Chapter 3 statistically explores the effects of systematic tonal correspondence between SC and JM vocabularies in speech production. This chapter focuses on how the strength of systematic correspondence is influenced by the sociolinguistic and cognitive backgrounds of the individuals. Between-word pitch distances of JM words can be predicted from SC tonal categories using statistical modeling, with interlingual tonal identity and individual backgrounds taken into consideration. The global influence of the bilinguals' sociolinguistic and cognitive backgrounds on the strength of systematic correspondence is found for the first time. Different from the other chapters, this chapter shows a more inter-disciplinary approach. The findings in this chapter are discussed in light of the theories of language change and bilingual cognitive aging. Their implications for speech engineering are also discussed.

Chapter 4 further investigates the effect of tonal similarity on etymologically related translation equivalents in auditory lexical access. The tonal similarity has discontinuous (facilitative or interfering) effects on the lexical access of SC-JM etymologically related translation equivalents. The tonal similarity also interacts with language mode (i.e. SC versus JM) and test blocks in a complex way. However, the persistence of the language dominance effect [i.e. With SC as the dominant dialect, the SC word is processed faster than its JM equivalent] indicates that the lexical representations of translation equivalents in the bilingual lexicon are not only distinguished by lexical tones but also by language mode. Moreover, the SC-JM bilinguals have an unusual lexical advantage compared with the tonal-monolingual controls, suggesting that bilinguals of closely related dialects, with their mental lexicon full of etymologically related translation equivalents, may benefit in their auditory lexical access.

JM also shows a considerable number of tonal lexical variants in speech production, which is investigated in Chapters 5 and 6. These two chapters try to answer how tonal lexical variants are stored and accessed in the bilingual lexicon.

Chapter 5 studies the role of different types of tonal variability in JM auditory lexical access, with a focus on the processing of tonal pattern variation between tonal lexical variants. True repetition, within-category variation, tonal pattern variation, and lexically contrastive variation are compared in an auditory form-priming experiment. The results support the view that tonal patterns have representative status in lexical access but also converge in a lexically specific way.

Chapter 6 compares JM lexical variants which are either identical or non-identical to their SC translation equivalents in lexical access for speech production. The variant probability effect suggests that the SC-JM bilinguals do store the JM tonal lexical variant they did not produce in the corpus, and that the difference in individual one-time choice still mainly reflects the variant probability instead of the individual difference in lexical representation.

Chapter 7 investigates the automatic bilingual phonological activation from the common written forms, with a focus on how the SC-JM tonal bilinguals differ from SC tonal monolinguals in retrieving tonal information from Chinese characters automatically. Although showing a general lexical disadvantage in visual word production, the bilinguals benefit more from the congruent conditions and suffer less from the incongruent conditions in this Stroop experiment. The SC-JM bilinguals are also different from the SC tonal monolinguals in their lack of tonal sensitivity, which may be related to the bilinguals' attention control and advantageous for the Stroop task.

At the end, Chapter 8 summarizes and discusses the main-findings of this thesis. The main-findings are restructured around the main theoretical issues. The role of systematic correspondence in language change and the impact of cognitive aging on tonal bilinguals' lexicon are discussed on the basis of the results from Chapter 3. The results from Chapters 2 and 4 are assessed together in the discussion of interlingual tonal similarity in bilingual lexical access. Similarly, Chapters 5 and 6 both investigate tonal lexical variants with tonal pattern variation. The results from these two chapters are discussed together, with a focus on the lexical storage of tonal lexical variants. Chapter 7 provides further insight into bilinguals' lexical access, but takes the automatic phonological processing in visual word recognition into consideration. And finally, bilinguals' attention and executive control, as well as their influences in bilinguals' lexical access, are discussed on the basis of the results of Chapters4 and 7.

中文摘要

本论文研究一种特殊的双语词汇加工系统。这个词汇加工系统涉及（1）两个声调系统，（2）两个密切相关的亲属方言，以及（3）一个通用的意音文字系统（logographic writing system）。

虽然声调在词汇通达（lexical access）中的地位得到了越来越多的关注，对于两个声调系统在双语心理词库（bilingual mental lexicon）中的相互作用，我们还是缺乏了解。并且，虽然历史语言学家对亲属方言词汇之间的系统对应（systematic correspondence）多有了解，却没有什么研究将系统对应现象与双语词汇产出（lexical production）中丰富的词汇变体与系统对应联系起来，过往研究也没有将系统对应与双语个体的社会及认知背景联系起来。另外，对于双语视觉词汇识别（visual word recognition），过往研究往往预设双语者要么使用两个相似的拼音文字系统（alphabetic writing system）要么使用两个不同类型的文字系统，却没有考虑到同一个意音文字系统也可以在两个密切相关的亲属方言之间通用。这种情况下，词源相关的翻译对等词（etymologically related translation equivalents）虽然在语音音系上有所不同，视觉上却使用完全相同的字形来书写。这个共同的字形如何激活双语者心理词库中的语音音系信息呢？这需要我们进一步的研究。

本研究以汉语普通话（Standard Chinese：SC）和济南方言（Jinan Mandarin：JM）作为研究案例，重新审视了一系列理论和实践问题，研究重点在于声调信息在词汇通达中的作用。汉语普通话是中国的官方语言，以北京话为语音基础。济南话是济南市的方言。两方言同属中国北方的官话方言。为了回答相应的研究问题，我在田野调查中通过行为实验搜集了一系列言语产生（speech production）和言语理解（speech comprehension）的数据。本文是一个论文集，因此具体的研究背景、讨论和结论都放在各个章节中。本摘要下文将对论文各章节进行简要介绍。

第一章为总引言。本章简要引入基本研究问题，并对相关研究领域进行综述。综述围绕本文所研究的双语现象的三个独特之处展开。首先，这个双语现象涉及两个声调系统。本章总结了声调语言在语音感知（speech perception）和词汇通达中的独特性，介绍了非母语声调感知学习（perceptual learning）及跨语言声调感知（interlingual tonal perception）的相关研究，并回顾了与汉语有关的双语研究。其次，这个双语现象涉及两个密切相关的亲属方言。不同于两个完全无关的语言，亲属方言之间既有相似的基本音位，又有词汇间的系统对应。例如，济南话单音节词的声调很大程度上能够从其普通话对等词的声调范畴推断出来。由于双语研究过去对系统对应不够重视，本章只对系统对应的机制及其在这个双语现象中的表现进行了集中介绍。第三，这个双语现象涉及一个通用的意音文字系统。本章首先回顾双语阅读中的跨语言激活，然后结合例子解释本文所研究的双语现象是如何使用文字的。对相关研究领域进行综述之后，我又分别介绍了正文各个章节的理论背景。

第二章研究这两个声调系统之间的二对一声调映射（two-to-one tonal mapping）。本章重点研究跨语言范畴优度（interlingual category-goodness）在

词汇通达和言语理解中的影响力。普通话的高升及低升调（第二声和第三声）与济南话的升调有二对一声调映射关系，这个关系在声学分布和跨语言感知上得到了证明。语义启动实验结果显示跨语言范畴优度仅影响词汇通达，但是对于词汇提取后的语义激活缺乏影响，这意味着这两个阶段之间存在一定的离散性。

第三章从统计上探索普通话与济南话声调的系统对应对言语产生的影响。本章重点研究双语个人的社会及认知背景如何影响系统对应的力度。统计建模发现，将跨语言声调同一性（interlingual tonal identity）和个体背景纳入考虑后，济南话两词之间的音高距离（between-word pitch distances）可以由普通话两词间的声调关系（tonal relations）预测。双语者的社会及认知背景对系统对应力度的影响是首次发现。与其他章节不同，本章展示了更为丰富的跨学科研究手段，研究成果的讨论同时引入了语言演变（language change）和双语认知老龄化（bilingual cognitive aging）的视角，并探讨了研究成果在言语工程（speech engineering）中的应用。

第四章进一步研究听觉词汇通达中声调相似性对词源相关的翻译对等词的影响。词汇通达中声调相似性在普通话—济南话对等词上的效应是非连续的（或为促进效应或为阻碍效应），并且声调相似性与普通话—济南话对等词的语言模式（language modes，如济南话或普通话）、以及测试区组（test blocks）都有复杂的交互作用。然而，语言优势效应（language dominance effect）稳定存在，这说明双语者心理词库中的词汇表征（lexical representation）不仅由词调（lexical tone）区分，也由语言模式区分。不仅如此，普通话—济南话双语者相对于声调单语对比组具有不寻常的词汇优势（lexical advantage）。这意味着，密切相关方言的双语者，或许是由于他们的心理词库充满了词源相关的翻译对等词，才能够在听觉词汇通达中获益。

济南话在语音产生中也呈现了不少声调词汇变体（tonal lexical variants），第五、六章对此展开研究。这两章试图回答双语心理词库中声调词汇变体是如何存储和通达的。

第五章研究济南话听觉词汇通达中不同类型声调变异的地位，重点探讨大脑如何加工声调词汇变体之间的调类变异（tonal pattern variation）。一个听觉形式启动实验比较了真实重复（true repetition）、范畴内变异（within-category variation）、调类变异（tonal pattern variation）以及词汇对立变异（lexically contrastive variability）。实验结果显示，调类在词汇通达中有表征地位，但是在此心理过程中，调类在也是可以汇合于具体单词的。

第六章在言语产生中比较两类济南话词汇变体，一类和其普通话对等词相同，一类和其普通话对等词不同。命名时间中的变体几率效应（variant probability effect）说明，即使在当前语料库中某个普通话—济南话双语者没有使用某个声调词汇变体，他依然存储了这个变体。个体对变体一次性的选择主要反映了变体的几率，而不是词汇心理表征上的个体差异。

第七章研究通用的字形如何在双语者身上产生语音激活，重点研究从汉字上自动激活声调信息的过程中，普通话—济南话的声调双语者与普通话的声调单语者有何不同。Stroop 实验条件下，双语者虽然在视觉词汇产生中总体呈

现了词汇劣势（lexical disadvantage），却不仅在字色一致的条件下获益更多，也在字色不一致的条件下损失较少。普通话—济南话双语者与声调单语者的另一点不同在于，他们在 Stroop 实验条件下对声调不敏感。这可能与双语者的注意力控制有关，并对 Stroop 任务有益。

最后，第八章总结并讨论了本文的主要发现，并根据主要理论问题重新组织了这些主要调查结果。基于第三章的实验结果，本章讨论了在语言变异和认知老龄化对声调双语者词库的双重影响下，系统对应的地位。而后，本章对第二章和第四章的实验结果进行综合考量，讨论了跨语言声调相似性在双语词汇通达中的地位。同样地，第五章和第六章都是研究涉及调类变体的声调词汇变体，这两章的实验结果放在一起，重点讨论声调词汇变体的词汇存储问题。第七章进一步讨论了双语者的词汇通达，但是将视觉词汇识别中的自动音系处理纳入考量。最后，本章综合第四章与第七章的结论讨论了双语者的注意力控制和执行控制，及其对双语者的词汇提取的影响。

Acknowledgement

First, I would like to express my gratitude to my supervisors, Yiya Chen, Niels O. Schiller, and Vincent J.J.P. van Heuven. The afternoon on which we decided the topic of this thesis is still fresh in my memory. Yiya wanted to see something about Chinese dialects, Vincent encouraged me to continue working on tones, Niels suggested that it would be a great idea to put bilingualism into perspective, and BINGO! The idea of tonal bilingualism came to mind. I would like to thank Yiya for her meticulous supervision throughout the whole process of my PhD. project and her help with my life in the Netherlands. I would like to express my gratitude for Vincent's editorial efforts in both English and Dutch and his introduction to the research environment in Leiden. I owe special thanks to Niels for his guidance in psycholinguistic methodology, the opportunity to practice my didactic skills, as well as his timely support at critical time intervals related to my defence. For all the challenges, guidance, inspiration, and support I have received from my supervisors throughout my studies in Leiden, I am very thankful.

I am grateful to the anonymous journal reviewers from *Language, Cognition and Neuroscience*, and *Speech Communication*, as well as my reading committee, Rint Sybesma (secretaris, Leiden), Claartje Levelt (Leiden), Ton Dijkstra (Nijmegen), Björn Köhnlein (Leiden), and Martijn Wieling (Groningen) for their careful reading, valuable comments and suggestions for the articles in my thesis.

My work was supported by a PhD Studentship sponsored by Talent and Training China-Netherlands Program, granted by China Scholarship Council (CSC) and The Netherlands Organisation for Scientific Research (NWO). The fieldtrips were sponsored by the Leiden University Foundation (LUF) and Leiden University Centre for Linguistics (LUCL). I would like to acknowledge all the support from these institutions.

There are many people who helped with my work in the various stages of this PhD. project. I would like to thank Prof. Xiufang Du, Prof. Jiangping Kong, Dr. Zihe Li, Dr. Honglin Cao, Jia Li, Lulu Zhou, and Dianliu Neighbourhood Committee for the recruitment of participants and providing spaces for the experiments. I would like to thank our lab engineer Jos J.A. Pacilly for his Praat scripts and all the technical advice. I would like to thank Dr. Martijn Wieling and Dr. Jacolien van Rij for their advice on statistics. I would like to thank Ailie Burningham and Mrs. Jenny de Sonneville for their English proofreading.

The Netherlands is a place where I met so many interesting and kind people. I would like to thank the more experienced Phd. students and Post-docs from LUCL, Franziska Scholz, Daan J. van de Velde, Jessie Nixon, Kalinka Timmer, Lesya Y. Ganushchak, Leticia Pablos, Willemijn Heeren, Jurrian Witteman, and Qian Li for their guidance in this amazing place. I would like to thank my Chinese friends (some of whom are also my colleagues) whom I met in Leiden and other parts of the Netherlands, Ting Zou, Qian Li, Man Wang, Lei Sun, Wenting Yu, Yifei Bi, Min Liu, Yangyang, Han Hu, Weixin Wang, Jinhua Wu, Kai Fang, Xiaofei Zhang, Shuo Mi, Hongyan Wang, Jing Lin, Jingwei Zhang, Liquan Liu, Yan Gu, Zenghui Liu, Xiaoli Dong, Anqi Yang, Lijie Zheng (Here I put their family names after their given names), as well as the teachers from the Confucius Institute of Leiden, for their inspirational insights into Chinese as well as the joyful times we spent together. I also would like to thank my language partner Björn Ooms, my housemates (‘de meisjes van Korte Hansjes’), my ex-landlady Jing Yang and her daughters Kitty and Cammy, my Dutch friends, Nienke Chin-van Oeveren, Joachim Bilars, Alain Corbeau, Bruce, Sietse Bärouwer, and Anna Kirstein (although she is actually a Dutch-speaking German) for integrating me into this beautiful culture. Special thanks to Cristian Bonato for his great dinner parties and outdoor activities, which brought us together through great music and original Italian food.

Last but not least, I would like to thank my parents for tolerating my delay of fulfilling the filial duties. I also owe my thanks to Tim Veldhuizen, especially for his patience, his insights as a proficient Scrum master, his editorial help, and his delicious ‘erwtensoep’, ‘draadjesvlees’, and banana cakes.

Curriculum Vitae

Junru Wu (吴君如) was born on August 29, 1985, in Yizheng, Jiangsu, China. At the age of four, she moved with her parents to Fuzhou, Fujian, China, and grew up there. She obtained a BA degree in Chinese Language and Literature from the Department of Chinese Language and Literature of Peking University in 2008 and a MA degree in Linguistics and Applied Linguistics (Phonetics) from the same department in 2011. From 2007 to 2011, she received extensive training in experimental phonetics in the Linguistic Lab of Peking University. During this phase, she participated in various lab- and field-works. She has explored a range of subjects, including sign languages, lip-reading, audio-visual speech processing, as well as speech phonation, and has published papers in several journals and conferences. She finally settled on tonal perception as the focus of her MA thesis.

Since the age of nineteen, she has been teaching part-time. Her teaching experience involves different target languages and students from various language communities, which draw her attention to language contact and bilingualism. To investigate a previously little studied type of bilingualism, tonal bilingualism of related dialects, in September 2011 she started her PhD research at Leiden University Centre for Linguistics, sponsored by the Talent & Training China program. This thesis is the result of this research project.